Search for $(W/Z \to jets) + \gamma$ Events in Proton-Antiproton Collisions at the Fermilab Tevatron

A thesis presented to the faculty of
The Rockefeller University
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

by

Andrea Bocci

2005

Abstract

We present a study of the $p\bar{p} \rightarrow W(Z)\gamma \rightarrow \gamma q\bar{q}$ process at the center-of-mass energy $\sqrt{s} = 1.96$ TeV using data collected by the Collider Detector at Fermilab. The analysis is based on the selection of low transverse momentum photons produced in association with at least two jets. A modification of an existing photon trigger was studied and implemented in the data acquisition system to enhance the sensitivity of this analysis. The data presented are from approximately 184 pb^{-1} of integrated luminosity collected by this new trigger. A preliminary event sample is obtained requiring a central photon with E_T >12 GeV and two jets with $E_T > 15$ GeV. The corresponding efficiency is studied using a Monte Carlo simulation of the $W(Z)\gamma \rightarrow \gamma q\bar{q}$ based on Standard Model predictions. Monte Carlo estimation of the background is not necessary as it is measured from the data. A more advanced selection based on a Neural Network method improves the signal-to-noise ratio from 1/333 to 1/71, and further optimization of the dijet mass search region increases the ratio to its final value of 1/41. No evidence of a $W/Z \rightarrow q\bar{q}$ peak in the dijet mass distribution is visible when the background contribution is subtracted. Using a fully Bayesian approach, the 95% confidence level upper limit on $\sigma(p\bar{p} \rightarrow W\gamma) \times \mathfrak{B}(W \rightarrow q\bar{q}) + \sigma(p\bar{p} \rightarrow Z\gamma) \times \mathfrak{B}(Z \rightarrow q\bar{q})$ is calculated to be 54 pb, which is consistent with the Standard Model prediction of 20.5 pb.

Acknowledgments

I would like to thank the entire Rockefeller University experimental high energy physics group and especially its unparalleled leader, Prof. Konstantin Goulianos who gives me the unique opportunity to do particle physics from the middle of Manhattan. His enthusiasm, knowledge and dedication to high energy physics (along with his endless anecdotes) were a true source of inspiration. I like to thank him for his continuous and unconditional support and to teach me that if you think hard enough you can always find a logical explanation to any result, even the most weird one.

Any graduate student should consider himself fortunate to have an outstanding advisor that is also considerate to give trenchant suggestions, patient in answering silly questions and always eager to provide guidance and support. Well, I was so lucky to have had three of such advisors: Anwar Bhatti, Steve Kuhlmann and Stefano Lami.

Stefano, my official advisor, always had my interest in mind and I am sincerely thankful for his advices and friendship. He has been a constant source of wisdom and understanding at any stage of this work (and even beyond) and not less precious have been his practical advices on how to cope with various aspects of the American society, from health insurance to credit history and a lot more. They made my life a little easier.

Steve is one of those rare physicists that every time you talk with them you wonder if there is something they are not expert on. This is impossible to express here how much I have learned from him. I think it is enough to say that for this work in particular all the ideas and the methods I developped in this analysis either were suggested by him or were conceived after a discussion with him. It was a truly privilege to work with him or under his supervision for all these years (since 1997 !) and as a physicist I could not ask for a better mentor.

I am very grateful to Anwar for his readiness to provide an answer to all my questions and concerns, every time with his typical relaxed attitude and never making me feel guilty for asking. What have impressed me the most about him however - I don't know if it is a gift or experience or both - is his incredible ability to grasp exactly my doubts even before I finished my sentence and not matter how bad I expressed them, and then promptly dispel them. His physics intuition will be always a reference for me.

I like to thank also: Luc Demortier for his super-condensed "Statistics for Dummies" discussion (perfectly tailored on me indeed) that allowed me finishing my

thesis on time; Christina Mesropian for sharing the office and thoughts and for putting up with me; Michele Gallinaro for his friendship and social support through the tides of my analysis progresses; Kenichi-san, our Kendo master, for his friendship and is invaluable help in handling the hideous tentacles of the CDF offline software; and the two Rockefeller group outposts at Fermilab, Mary Convery and Koji Terashi.
Finally I also thank all the people and friends I met at Fermilab to have made my visits to Batavia bearable, and sometimes even enjoyable.
And then there are all the people outside the field I met here in New York to thank, but I won't do that. They would not read those lines anyway.
This thesis would not have been possible without the dedicated effort of the entire CDF collaboration. All CDF members have my sincerest thanks.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

There have been profound developments in particle physics in the past quarter century, comparable with many important discoveries through the history of physics. The minimal requirement for claiming that (a segment of) nature is understood is the particles and the forces acting on them, and the rules for calculating the effect of the forces. The last of these came first, with the discovery of the quantum theory. The evidence is increasingly better that the Standard Model, in which gauge forces mediated by gauge bosons act on quarks and leptons, describes nature's basic behavior. The discovery of the top quark in 1994 and the direct observation of the tau neutrino in 2000 at the Fermi National Accelerator Laboratory are two of the most recent experimentally verified Standard Model predictions.

Despite this phenomenal success, however, there is a whole segment of the theory that has not been observed: the Higgs sector. If the Higgs particle turns out to exist as conventionally described, then that closes the Standard Model from a mathematical point of view. Moreover, uncovering the secrets of the Higgs sector might provide essential clues to further understanding of matter. Why all know particles acquire their masses exclusively through the Higgs system (the theory does not require that), the relationship between the Higgs sector and the cosmological constant, or how the Higgs boson can distinguish among fermions of different families (coupling to the muon 200 times stronger than to the electron, for example) are just few of the particle physics mysteries related to this sector of the Standard Model.

It is not a surprise, therefore, that the search for the Higgs boson occupies the center-stage of the present and future experimental research. The Tevatron is currently the only machine in the world able to probe Higgs sector. It has been shown (see sec. 2.2) that the sensitivity of such a search depends critically upon attaining the best possible resolution of the mass measurement of the two jets coming from the Higgs decays.

The identification and reconstruction of the jets coming from a boson hadronic decay is the subject of this dissertation. Given the very small cross section expected for the processes involving the Higgs, a direct search appears to be

not feasible at this stage of the Tevatron Run II. As a matter of fact, not even some of the main background processes (like the diboson semileptonic decays $W(W/Z) \to l(\nu/l)q\bar{q}$) have been observed yet, although their cross sections are almost two orders of magnitude higher than the Higgs one.
The main difficulty in identifying such processes arises from the huge QCD multi-jet background a dijet boson resonance has to be extracted from. In general, all searches for dijet resonances at hadron colliders (whether they are Standard Model particles as Higgs, W, Z, or beyond the Standard Model particles as W', Z', Randall-Sundrum gravitons, axigluons, E_6 diquark,...) suffer from this kind of problem whose importance increases with the size of the cross section of the processes involved.
The analysis reported in this dissertation addresses this issue by looking for the hadronic decay of the W and Z bosons in photon events. Although the difficulties in extracting the $W/Z \to jets$ signal in such events are very similar with those related to the aforementioned W(W/Z) semileptonic decay, the hadronic decays have the advantage of a ten times larger cross section. This implies that the chances to detect a $W/Z \to jets$ signal are much higher in photon events than in any other class of processes where these bosons are produced along with a large QCD background [1]. As we mentioned already this is not an easy task. In fact not only in Run I but even in Run II the W/Z hadronic resonances have not been observed so far, in spite of the four times larger amount of data available. Therefore it is very important to show that a mass resonance signal is extractable when it sits on a large QCD background, and it makes sense to do that using the data sample that gives the highest chance of success. For these reasons the focus of this dissertation is to investigate the best techniques and tools necessary to discriminate the dijet signal from the generic background jets and apply them to the search for the $W\gamma$ and $Z\gamma$ events.
In addition, the diboson production in which the W and Z are produced in association with a photon is interesting in its own right. These events have been used to test the structure of the non-Abelian character of the Electroweak theory. Although the leptonic decay channels of the bosons are easier to detect, an independent study in the hadronic channels can provide a confirmation of their results and eventually the physics outcomes can be combined to further improve the accuracy of the overall measurement. Moreover, once the W/Z peak is established, exotics searches can be undertaken using these events. There are some extensions of the Standard Model that predict the existence of more gauge bosons, called W' and Z', which decay predominantly into ordinary W and Z and a photon. Other theories instead predict a different value for the $WW\gamma$ tri-linear boson coupling and a different from zero $ZZ\gamma$ and $Z\gamma\gamma$ couplings. All these modifications of the Standard Model can be profitably tested studying the

[1] It is important to notice that there are processes where the W dijet resonance can be successfully reconstructed because of a low background contamination (as for example the $W \to q\bar{q}$ peak in top events (sec. 2.3.2). However this is not the spirit of our search. Our goal is to identify W/Z hadronic resonances *extracting* them from the QCD background, in order to learn how to do that and possibly to establish a method that can be applied to similar searches.

$W\gamma$ and $Z\gamma$ events.

1.1 Overview of the Thesis

This dissertation is structured as follow:

In Chapter 2 we outline the main feature of the Standard Model and we provide an historical review of the hadronic resonance searches at hadron colliders.

In Chapter 3 we describe the experimental apparatus, the parameters of the detector components and their performance. A brief description of the data acquisition system is also provided.

In Chapter 4 we review the methods and the algorithms commonly used at CDF to reconstruct the physics objects. A particular attention is devoted to those objects relevant to this analysis, photons and jets.

In Chapter 5 we introduce the Monte Carlo sample used to simulate the $W(Z)\gamma \rightarrow \gamma q\bar{q}$ signal events and to study the signal characteristics. The expected $W(Z)\gamma$ cross section is also calculated based upon Standard Model predictions.

In Chapter 6 we describe the new dedicated trigger we implemented in the CDF data acquisition system for this analysis and its efficiency for the identification of the $W(Z)\gamma \rightarrow \gamma q\bar{q}$ events.

In Chapter 7 we present the selection criteria applied to our dataset and we discuss their acceptance and efficiency along with the corresponding uncertainties.

In Chapter 8 we compare the $\gamma + 2$ *jets* background Monte Carlo sample with the data and we discuss the accuracy of the simulation in reproducing the background of our analysis.

In Chapter 9 we present the neural network advanced event selection introduced to optimize the signal-over-background ratio of our sample. We show how the most discriminating kinematic variables are selected and the effectiveness of this approach and its efficiency. At the end of the chapter we measure the expected background contribution on the final selected data and the $W/Z \rightarrow q\bar{q}$ peak search is performed.

In Chapter 10 we calculate the 95% confident level upper limit for the $W(Z)\gamma$ cross section times the BR($W/Z \rightarrow q\bar{q}$) using a full Bayesian approach and comparing the result with the Standard Model prediction. We then discuss the future prospects of this search in Run II.

In Chapter 11 we present the summary and the conclusions of the analysis described in this dissertation.

Chapter 2

Standard Model and Present Understanding

2.1 The Standard Model

The foundation of the contemporary understanding of the fundamental constituents of matter and their interactions is expressed in a theory called the Standard Model. The theory bases were developed in the 1960's and 70's and has been extensively tested experimentally. When predictions can be made, the experimental agreement with the theory has been excellent.
The Standard Model integrates two separate theories: the Quantum Chromodynamics (QCD) [2], describing the strong interaction, and the Glashow-Weingberg-Salam (GWS) theory [1], which unifies the weak and the electromagnetic interactions. They are described using the theoretical framework of the quantum field theory, and therefore the Standard Model is consistent with both quantum mechanics and special relativity.

The Elements of Matter

All the fundamental particles can be classified into two groups according to their spins. These are the *fermions*, which are spin-$\frac{1}{2}$ particles and the *bosons*, which are spin-0 or spin-1. Fermions are split in two groups, *quarks* and *leptons*, and both are arranged in three families:

$$\begin{pmatrix} u \\ d \end{pmatrix} \quad \begin{pmatrix} c \\ s \end{pmatrix} \quad \begin{pmatrix} t \\ b \end{pmatrix} \quad \text{quarks}$$
$$\begin{pmatrix} e \\ \nu_e \end{pmatrix} \quad \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix} \quad \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix} \quad \text{leptons}$$

The electron e, the muon μ and the tau τ are massive, have negative electric charge (-1), and interact both electromagnetically with their antiparticle counterparts e^+, μ^+ and τ^+, and weakly with their corresponding neutrinos

ν_e, ν_μ and ν_τ. Being electrically neutral, the neutrinos experience only weak interactions.
Quarks come in six varieties, known as flavors, called up (u), down (d), charm (c), strange (s), top (t), and bottom (b). Along with carrying an electric charge, which is a precise fraction of an electron charge, all quarks have "color" charge (red, blue or green), the charge of the strong interaction. Quarks are *confined* in groups with other quarks to form *hadrons*, colorless composite particles. Hadrons are divided into *baryons* (three quark objects), and *mesons* (quark-antiquark objects). Ordinary matter is made of baryons such as protons and neutrons.
The other group of fundamental particles are the carriers, or mediators, of the interactions, known as the gauge bosons. The carrier of the electromagnetic interaction is the photon. The weak interaction is mediated by three vector bosons, W^+, W^- and Z^0, and gluons are the carriers of the strong interactions. They carry color charge in color-anticolor pairs. The strength of each force is given by its coupling constant, α, and the average distance over which it acts is given by its range.

2.1.1 The Electroweak Theory

The Electroweak theory [1] is a unified description of the electromagnetic and weak forces. It combines the weak isospin group (non-Abelian) SU(2) and weak hypercharge group (Abelian) U(1) in SU(2)×U(1). The weak force distinguishes between left and right handed [1] components of fermions and allows parity (P) and charge conjugation (C) violations in weak processes. The usual notation for the electroweak gauge group is $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$. The subscript L stands for the left-handed fermions and Y for hypercharge. Before the electroweak unification, the electromagnetic group has a $\mathrm{U}_Q(1)$ symmetry generated by the electric charge Q and the weak interaction has a $\mathrm{SU}_L(2)$ symmetry from the charged and neutral current interaction. In order to preserve the $\mathrm{SU}_L(2)$ symmetry when constructing the isospin triplet of weak currents, it becomes necessary to modify the U(1) electromagnetic group generator to account for the right-handed interactions. The hypercharge Y is then introduced and replaces electric charge as a group generator with the definition $\mathrm{Y}=2(\mathrm{Q}+\mathrm{T}_3)$, where Q is the electric charge and T_3 the third component of the weak isospin of the particle. The electroweak theory is built around the conservation of the weak isospin and weak hypercharge making the Lagrangian invariant under local gauge transformations. The gauge invariance of the electroweak Lagrangian is though complicated by the mass of the carriers of the weak force ($\mathrm{W}^\pm$ and Z). Therefore, besides the gauge invariance, the mechanism called spontaneous symmetry breaking needs to be introduced. It gives rise to the massive gauge bosons and to a particle that holds the key to the origin of mass: the Higgs boson. It is responsible to generate the large masses for the weak gauge bosons ($m_W \sim 80$ GeV, $m_Z \sim 91$ GeV) explaining the short range of the weak force.

[1] Massless right handed particles have the component of the spin in the same direction of motion, the left handed ones in the opposite direction.

2.1.2 Quantum Chromodynamics

Quantum Chromodynamics (QCD) [2] is a renormalizable theory modeled after the Quantum Electrodynamics (QED), the gauge theory of the electromagnetism. In QED the $U_Q(1)$ symmetry requires the electric charge to be conserved. Similarly, the QCD SU(3) symmetry requires the *color* charge to be conserved. However, unlike QED, the gauge symmetry is non-Abelian, causing gluons also to posses color charge and consequently interact with themselves as well as with quarks. Moreover, the additional gluon-gluon interactions cause the strong coupling constant α_s to have a qualitatively different behavior with Q^2 (the interaction momentum transfer scale) than the QED coupling constant α_{QED}. To a first approximation in Q^2/Λ_{QCD} one has:

$$\alpha_s(Q^2) = \frac{4\pi}{(11 - \frac{1}{2}N_f)\ln\frac{Q^2}{\Lambda_{QCD}}}$$

where N_f is the number of quark flavors with mass less than Q and Λ_{QCD} is a parameter which, qualitatively, indicates magnitude of the scale at which $\alpha_s(Q^2)$ becomes strong. Λ_{QCD} is determined experimentally to be about 0.2 GeV. The strong coupling constant then becomes small for a large transfer momentum , and a perturabitve description of the strong force is possible. However, at momentum transfer comparable with the masses of the light hadron (*i.e.* $Q^2 \simeq$ 1 GeV), α_s becomes large and perturbation theory breaks down. This large value of the coupling constant is the source of most of the mathematical complexities and uncertainties that still surround QCD calculations at low Q^2. On the other hand it is of great importance that this "running" coupling goes to zero in the infinite Q^2 limit. This fact, called *asymptotic freedom*, allows perturbation theory to be used in theoretical calculations to produce experimentally verifiable predictions for hard scattering processes.

2.1.3 QCD Scattering Formalism

This thesis deals with physics processes produced in proton-antiproton collisions. Since the proton and antiproton are composite particles, in order to be able to theoretically calculate the outcome of a $p\bar{p}$ collision, their internal structure must be considered.
The proton consists essentially of three quarks (uud) - called *valence* quarks - that establish its charge and other quantum numbers. In addition to them, a fluctuating cloud (or *sea*) of virtual gluons and neutral $q\bar{q}$ pairs is expected to be present in each hadron. Quarks and gluons inside a hadron are generally referred as *partons*. Thus, the proton and antiproton can be seen as "broad-band" beams of partons carrying varying fractions of the momentum of their parent hadron.
The $p\bar{p}$ collision can be pictured as a sequence of three different phases occurring at different time scales (fig. 2.1):

- partons approach each other with some momentum distribution – "Parton Distribution Function, PDF" – inside the parent hadron.
- A hard collision takes place between a pair of partons regarded as free particles. The cross section can be predicted by perturbation theory.
- New partons are generated by the two scattered partons and subsequently the quarks and gluons rearrange themselves into hadrons. This process is called *hadronization* or *fragmentation*.

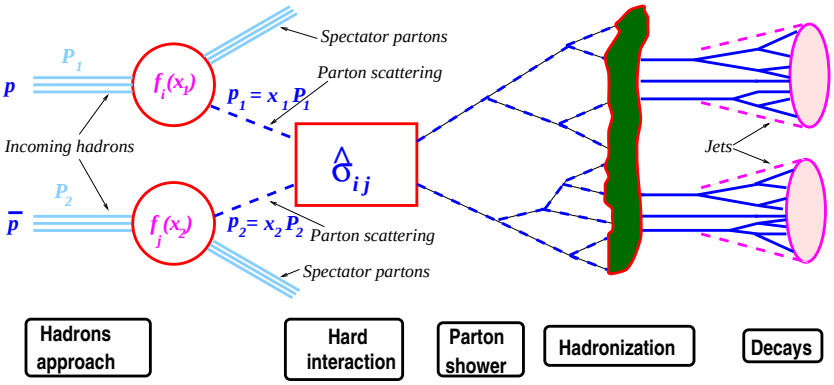


Figure 2.1: *Sketch of a hard proton-antiproton interaction. P_1 and P_2 are the momenta of the incident hadrons, with $p_1 = x_1 P_1$ and $p_2 = x_2 P_2$ being the momenta of the partons participating in the hard interaction. The cross section for the scattering of the partons of types* i *and* j *are denoted by $\hat{\sigma}_{i,j}$ and it is calculated perturbatively.*

The contributions from soft interactions which occur long before the hard scattering[2] can be "factored out" and absorbed into the effective momentum spectrum of the incoming hadrons, typically in the Parton Distribution Functions. In this way the remaining cross section involves only high momentum transfers (and therefore short times and distances) and in particular does not depend on the details of the hadron wave function or the type of hadrons. It is a single short-distance interaction and is computable in perturbation theory thanks to asymptotic freedom. After the scattering, partons lose their energy in a perturbative evolution to a lower virtual mass ($\sim$ 1 GeV), thorough gluon bremsstrahlung and $q\bar{q}$ pair production ("*parton shower*"). Hadronization occurs at a much later time scale characterized by $1/\Lambda_{QCD}$, where Λ_{QCD} is the

[2]A very important theoretical issue is whether the partons in hadron H_1, through the influence of their color fields, change the distribution of parton in hadron H_2 before the hard scattering approximation is applied, thus spoiling the simple parton picture that we have outlined. Soft gluons which are emitted long before the collision are potentially troublesome in this respect. This is a feature not present in process involving only one incoming hadron (as Deep Inelastic Scattering) but it is distinctive of hadron-hadron interactions. The *theorem of factorization* [3] states that since the initial and final distributions of partons interact on time scales that are vastly different from that of the hard scattering, the interference between the hard scattering and the initial and final states should be small. The property of factorization allows us to use the QCD parton model to describe the inelastic process.

scale in α_s at which the coupling becomes strong. The interactions which change quarks and gluons into hadrons certainly modify the outgoing state, but they occur too late to modify the original probability for the scattering events to happen, which can therefore be calculated in perturbation theory.

2.2 The Higgs Search

Despite the phenomenal success of the predictive powers of the Standard Model, there is a portion of the framework for which no evidence has been discovered. We already mentioned that the local gauge symmetry of the electroweak interactions fails to predict massive carriers for the weak interactions. To correct this flaw of the theory, a spontaneous symmetry breaking was introduced, initiated by what is called the Higgs mechanism [5]. Moreover it is through the interaction between the Higgs field and the fundamental fermion fields that the masses of the latter arise.
However there is another critical consequence of the Higgs mechanism that still has not been confirmed experimentally: the existence of a neutral scalar boson generated by the self interaction of the Higgs field and called just *Higgs boson*. The mass of the Higgs is not predicted by the theory and it needs to be measured experimentally. Since its value is deeply connected to the breaking scale Λ - the energy scale the Standard Model must break down and new physics must come into play - a discovery of the Higgs not only offers insight into the electroweak spontaneous symmetry breaking mechanism, but also sheds light on the ultimate reach of the Standard Model relevance.
Exploiting the interdependence of the parameters of the theory (in particularly between the mass of the W, the mass of the quark top and the mass of the Higgs) an upper limit for the Standard Model Higgs mass has been determined. Combining this upper limit with a lower limit coming from a direct search performed by the LEP2 experiments, a mass range between 114-260 GeV [6] [3] is set for the Higgs mass [4].
After the LEP machine was shut down in 2000 and before the Large Hadron Collider turns on, the Tevatron is the only machine in the world capable of probing the Higgs mechanism. Maximizing the sensitivity of the search for the Higgs - that for masses within the Tevatron range decays predominantly into $b\bar{b}$ pairs - depends most critically on attaining the best possible $b\bar{b}$ mass resolution (fig. 2.2). In the plot in fig. 2.3 the expectation for a Higgs discovery/limit with the new Run II detector acceptance and an improved di-jet mass resolution (compared to that in Run I) is showed [7].

In order to address the crucial request of an improved dijet mass resolution, a very general study on jet algorithms was carried out at CDF [9], resulting in a new jet energy reconstruction method based on the event particle energy

[3] The Higgs mass upper limit depends strongly upon both the mass of the top and the W used in the computation. The 260 GeV reported here was deduced assuming $M_W = 80.41 \pm 0.4$ GeV and $M_t = 178.0 \pm 4.3$ GeV

[4] Here we refer always to the Standard Model Higgs.

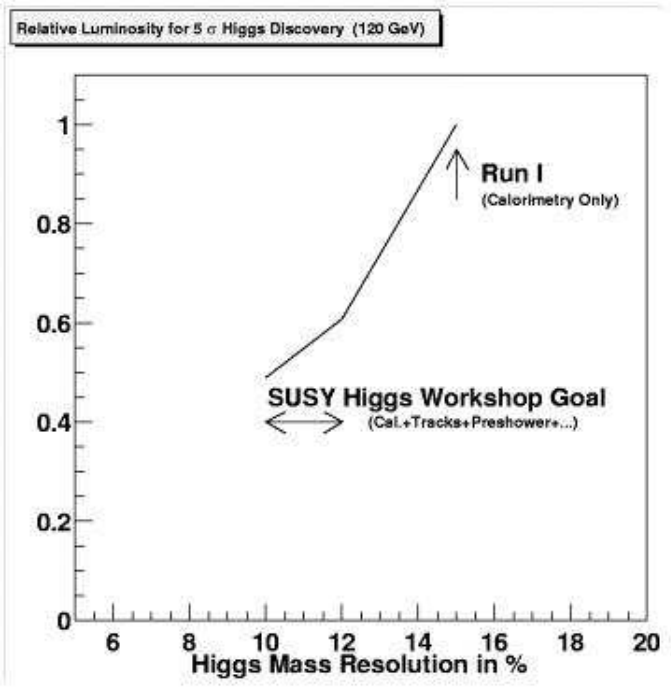


Figure 2.2: *The relative luminosity needed for a 5-sigma discovery as a function of the Higgs mass resolution. The value used by the Higgs Working Group [7] is also shown. The mass of the Higgs is assumed to be 120 GeV.*

flow. In spite of the jet resolution improvement showed by applying this method to "γ+jet" data events, it was not possible to test it on dijet events of know resonances (like W and Z) and evaluate its true potential, essentially because CDF was *never* able to isolate the hadronic decays of the weak bosons, W and Z.
Since it is the main subject of this dissertation, in the following paragraphs we outline the various searches for the hadronic decays of the W and Z performed in the past at hadron colliders.

2.3 W/Z Hadronic Decays at Hadron Colliders

The W and Z bosons were discovered at CERN in 1983 [4] providing the most convincing evidence for the electroweak theory. Their masses and decay characteristics are very well estimated by the theory with an impressive agreement with the experimental values. The W and Z bosons can be detected using their product decays, a pair of fermion-antifermion ($f\bar{f}'$). The fraction of the W(Z) decaying in a particular pair - or channel - $f\bar{f}'$ is called *Branching Ratio* of the

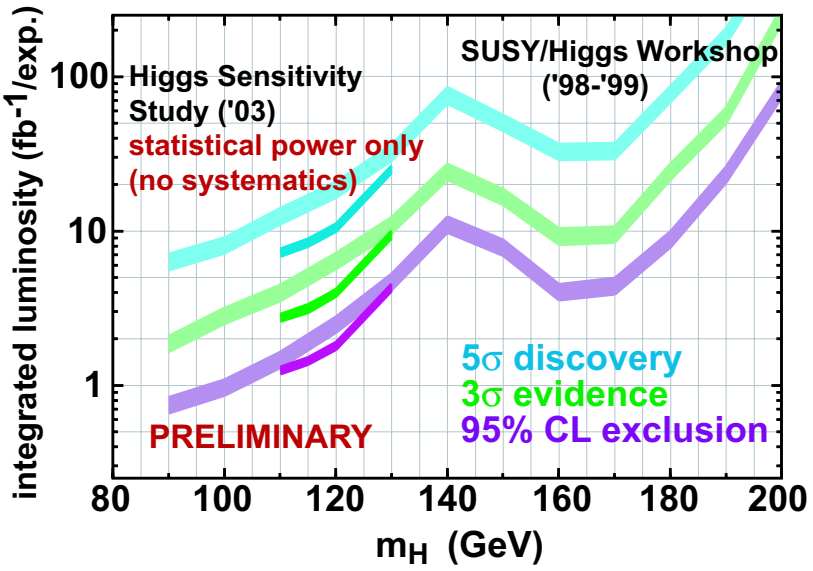


Figure 2.3: *The luminosity threshold for the 95 % CL, 3σ excess, and 5σ excess as a function of standard model Higgs mass. This is for the combined WH and ZH channels for both experiments. The width of the bands from this study is approximately 10% and represents the statistical errors.*

channel $f\bar{f}'$ and it is indicated as $\mathfrak{B}(W(Z) \to f\bar{f}')$. Table 2.1 shows all the possible decay channels for both the W and Z.

Intermediate Vector Boson Decays		
Process	Relative Coupling	Branching Ratio
$W^+ \to e^+\bar{\nu}_e, \mu^+\bar{\nu}_\mu, \tau^+\bar{\nu}_\tau$	1	$3 \times 11.1\%$
$W^+ \to u\bar{d}$	3	33%
$W^+ \to c\bar{s}$	3	33%
$Z^0 \to \nu_e\bar{\nu}_e, \nu_\mu\bar{\nu}_\mu, \nu_\tau\bar{\nu}_\tau$	$\frac{1}{2}$	$3 \times 6.8\%$
$Z^0 \to e^+e^-, \mu^+\mu^-, \tau^+\tau^-$	$\frac{1}{4} + (\frac{1}{2} - 2\sin^2\theta_W)^2$	$3 \times 3.4\%$
$Z^0 \to u\bar{u}, c\bar{c}$	$3[\frac{1}{4} + (\frac{1}{2} - \frac{4}{3}\sin^2\theta_W)^2]$	$2 \times 11.8\%$
$Z^0 \to d\bar{d}, s\bar{s}, b\bar{b}$	$3[\frac{1}{4} + (\frac{1}{2} - \frac{2}{3}\sin^2\theta_W)^2]$	$3 \times 15.2\%$

Table 2.1: *Vector boson decay into fermion-antifermion. The θ_W is the Weinberg angle ($\cos\theta_W \approx M_W/M_Z$).*

The quark decays have an higher branching ratio (because of the color factor) and therefore both the W and Z bosons prefer to decay into two jets. However

these decay channels are very difficult to observe in $p\bar{p}$ collisions due to the large QCD dijet background. In fact, although the W and Z have been discovered in a hadron machine, the precise measurement of their properties were obtained in e^+e^- collider experiments. In $p\bar{p}$ machines they are reconstructed exclusively through their leptonic decays, which have a clear signature and a very small background.
As we will see next, in spite of the very difficult environment, there were in the past some attempts to reconstruct the W and Z bosons in the hadronic channels.

2.3.1 The UA2 results

The first observation of the W and Z boson hadronic decays was reported by the UA2 experiment [8]. Using about 4.2 pb^{-1} of data collected at the CERN $p\bar{p}$ collider with a $\sqrt{s}$=630 GeV, a signal was reported in the dijet invariant mass spectrum, above the smooth QCD background (fig. 2.4). The readout system

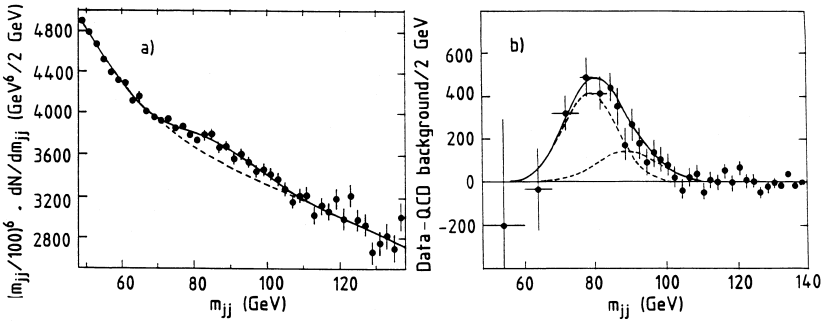
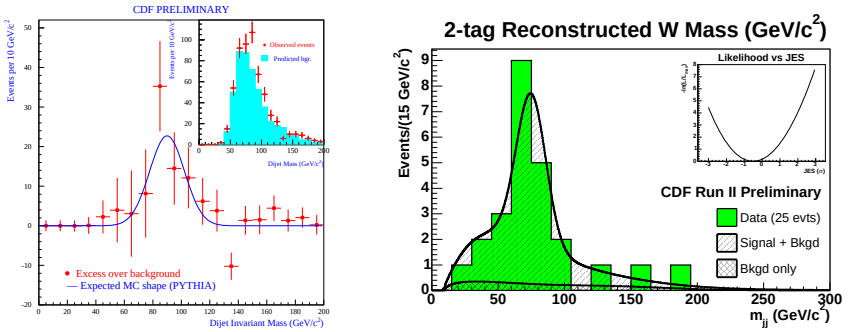


Figure 2.4: *The signal $W/Z \rightarrow$ jets seen by UA2. (left) The dijet mass spectrum in the region 48< m_{jj} <138 GeV. A combined fit to QCD background and W/Z signal is shown (full line) along with the background contribution only (dashed line). (right) Dijet mass spectrum after the subtraction of the fitted QCD background. The fit is performed with a double Gaussian with the resolution left as a free parameter.*

was specifically modified for this measurement to collect as much amount of dijet events as possible. Special care was taken to have a *large* low-mass control region ($m < 65$ GeV, where no signal contribution is expected) to strictly constrain the extrapolation of its shape inside the signal region. In fact it was shown that the significance of the signal observed would have halved if the lower mass threshold were 51 GeV instead of 45 GeV and no significant signal would have been found if the lowest mass threshold had been as high as 57 GeV, as a consequence of a small size control region. A high mass control region ($m > 105$ GeV) was used as well to lock the background shape as it bridges across the signal region toward higher masses.
A function of the form $f(m) = m^{-a}e^{-bm}e^{-cm^2}$ was used to fit the dijet invariant mass spectrum excluding the range 70< m_{jj} <100 GeV (*i.e.* the signal region). The best fit produced a signal of 5618±1334 events, corresponding to a signal



(a) $Z \to b\bar{b}$ (b) $W \to q\bar{q}$

Figure 2.5: *(a) $Z \to b\bar{b}$ peak extracted from high P_T muon event in Run I; (b) Run II $W \to q\bar{q}$ in $t\bar{t}$ events where two other jets have been identified as b-jets ("2-tag") and the other W decays leptonically.*

to background ratio of 1/38.
In this dissertation we will adopt a similar technique to search for the W/Z hadronic peak, extrapolating the background from the two control regions at low and high mass inside the signal region. As we will describe in details, a lot of effort was spent during the selection of the data sample in order to have an as large as possible low mass control region, a condition that the UA2 experiment proved essential for the success of this method.

2.3.2 Two Jet Hadronic Resonance Search at CDF

If extracting the W/Z peak from the dijet QCD background was rather challenging for UA2, moving from $\sqrt{s}$ =630 GeV of the $S\bar{p}pS$ collider to the $\sqrt{s}$ =1800 GeV of the Tevatron made this kind of search virtually impossible. At this energy the QCD dijet production is so high that even one hundredth of such an events is an amount that would exceed the entire experiment storage capability. To cope with this huge rate a relatively high threshold (20 GeV) is applied to the jets and only one out of 1250 of such a events are recorded [5]. Therefore at CDF it was not possible in Run I (let alone in Run II) to search for the W/Z hadronic decays in the general way as the UA2 experiment did. The only serious attempt was done at the end of Run I (Run IC) when a field wire in the tracking system broke requiring to turn off part of that detector. As a consequence the dijet threshold was lowered and more bandwidth became available. However, even under these very favorable circumstances, the lack of a low dijetmass control region - along with the small S/N and the insufficient knowledge of the p_t threshold bias - did not allow a successful identification of the W/Z peak [10].

[5] To make things even worse, for this energy threshold the events are not even full efficient for a W/Z signal search.

Ruling out a direct search in dijet events, CDF was able to identify the W and Z hadronic decay in two very special cases.
Using the high P_T muon data sample a $Z \to b\bar{b}$ signal was extracted [11] in Run I (and with a similar method it was established in Run II as well). The selection of only b-jets drastically reduces the QCD background, increasing the S/B to an acceptable level. Moreover it allows the determination of the background spectrum from an independent sample, waiving the need to have a control region in the dijetmass spectrum itself (fig. 2.5(a)).
A $W \to q\bar{q}$ signal was instead extracted [12] using $t\bar{t}$ candidates in the decay mode $t\bar{t} \to (bW^+)(\bar{b}W^-) \to b\bar{b}q\bar{q}l\nu$. Since the two b-jets can be identified, the other two jets must be the decay product of one of the W (the other one decay leptonically). The reconstruction of these events is very clean because of the (almost) total absence of background (fig. 2.5(b)). Unfortunately the number of such an events is very limited, even if a significant amount of them are expected to be collected during the Run II.
Besides these two successful examples, other searches for the W/Z peak have been performed (with no avail so far) using diboson events, as we discuss next.

2.3.3 $W(Z)\gamma \to \gamma q\bar{q}$ Search in Run I

Another class of processes where the W/Z hadronic decays can be searched is the diboson production [6] (WW, WZ, $W\gamma$ and $Z\gamma$). Although the cross sections of such processes are much smaller than the W/Z direct production, the presence of another object strongly suppresses the QCD background making these events very appealing.
The limited statistics available so far did not allow the identification of these processes yet, but with the data CDF has been collecting in Run II, the chance to finally observe diboson events in the semileptonic channels becomes possible. The analysis reported in this dissertation indeed is one of the ongoing efforts made in this direction. The $W(Z)\gamma$ production has the great advantage of having a cross section significantly higher than those for heavy diboson production and, therefore, its experimental observation should require a smaller amount of data and be achieved sooner.
An analysis in the same channel was already carried out in Run I [13]. Starting from a sample of events with a photon of at least 25 GeV, the requirement of two reconstructed jets were added. Then a search for the ~27 events from the W/Z hadronic decay expected in the sample was performed. The large signal-over-background ratio (~1/100) along with the small amount of signal events expected, did not allow the identification of the W/Z peak over the huge QCD background. However this study provided useful indication on how to undertake this analysis in Run II.

[6] In the case of two heavy diboson, one of them is required to decay leptonically in order to be easily identified (this configuration is called semileptonic). In fact if both bosons decay hadronically, their identification from the four jets final state is even more difficult that the direct search.

2.4 $W(Z)\gamma \to \gamma q\bar{q}$ Search in Run II: Strategy

Before we embarked in the challenging search of the W/Z hadronic peak in Run II photon events, we tried to bring forth a strategy to enhance our chance of success summarizing the lessons learned from the past studies outlined above in these three main analysis guidelines:

1. Collect as many signal events as possible, in particular lowering the photon energy threshold and increasing the overall acceptance.
2. Provide a dijet mass spectrum with a low mass control region to constrain the background fit and extrapolate it inside the signal region (*à la* UA2)
3. Maximize the signal-over-background ratio exploiting all the possible differences between events having jets coming from the heavy boson decays and those having just generic jets coming from QCD.

As we will describe in the following chapters, throughout this analysis we always try (in order to get the most out of the data at our disposal) to comply to these general rules.

Chapter 3

Experimental Apparatus

3.1 Detector Overview

The Collider Detector at Fermilab (CDF) started to be operational in 1985 and shortly after the first proton antiproton collisions produced in the Tevatron accelerator were detected by the CDF detector. Since then, through several period of data taking (see Table 3.1), it collected the largest amount of $p\bar{p}$ data than any other experiment at the highest energy ever achieved in an accelerator machine.

A plethora of outstanding physics outcomes - as the first evidence of the existence of the top quark - have been produced from the analysis of the data collected during the so called Run I (1992-96) resulting in more than 200 publications.

Run	**Period**	**Integrated Luminosity (pb$^-$1)**
-	1987	0.025
Run 0	1988-1989	4.5
Run IA	1992-1993	$\sim$ 19
Run IB	1994-1995	$\sim$ 90
Run IC	1995-1996	$\sim$ 1.9
Run IIA (08/2004)	2001-2004	$\sim$ 400

Table 3.1: *CDF $p\bar{p}$ data taking history*

After the end of Run I in 1996, in order to further enhance the capability of the machine, Fermilab had undergone a major upgrade of the Tevatron accelerator complex increasing both the center-of-mass energy of the proton-antiproton collisions (up to $\sqrt{s} = 2$ TeV) and the luminosity of the beam (see Appendix A). As a consequence of the new improved running condition of the accelerator, in order to accomodate the higher collision rate, most of the existing CDF components have been rebuilt or upgraded, providing improved capability as

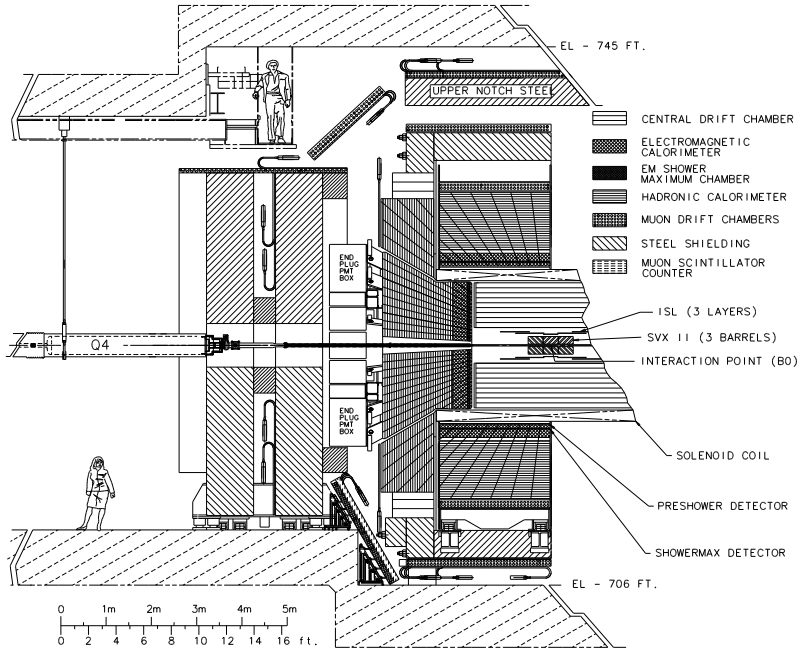


Figure 3.1: *An elevation view of the Collider Detector at Fermilab in Run II*

well as extended coverage with respect to the Run I counterpart. The tracking system has been totally replaced with new detectors. The calorimetry system is now exclusively scintillator-based. The electronics and trigger systems are fully compliant with the new pipeline configuration, and all software has been re-written using C++ and an Object Oriented architecture. The elevation view of the upgraded detector (CDFII) is shown in fig 3.1.

The CDF layout is standard for this type of detectors. From the inside out there are: tracking system, magnet, electromagnetic calorimetry, hadron calorimetry, and muon detectors on the far outside. There are also systems, such Time-of Flight, Cherenkov Luminosity Counter and Forward Detectors, which are not as universal as those just mentioned, but still are extremely important for a particular class of analysis. The overall geometry of the detector is cylindrical, with the Tevatron beam-line running through the center, along the axis of symmetry. CDF employs a right-hand global coordinate system with the origin in the very center of the detector (the nominal interaction point, IP). It is defined with the clockwise travel of the protons being the positive z direction with the y axis points upward (see fig 3.2).

The azimuthal angle ϕ runs in the transverse $(x-y)$ plane, with $\phi = 0$ being the positive direction of the x axis. Polar angle θ is counted from the positive direction of the z axis. Often, the polar angle coordinate θ is replaced by the

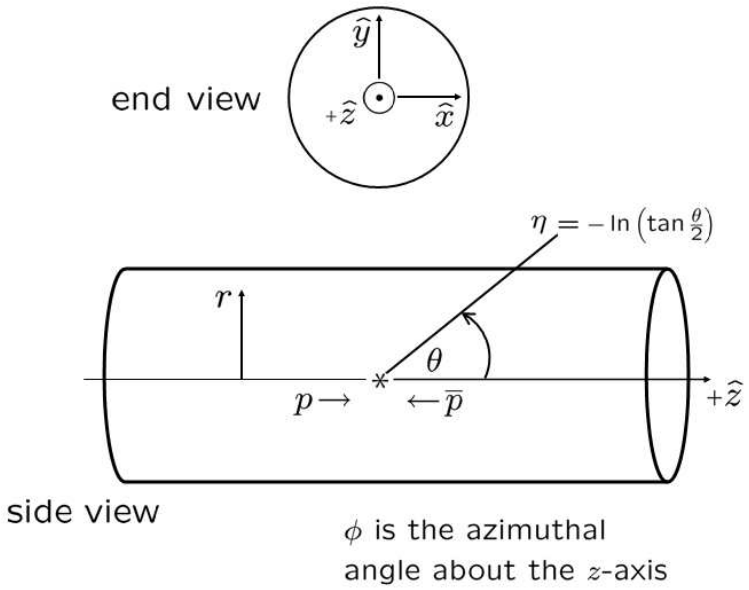


Figure 3.2: *Definition of the CDF coordinates*

quantity (called *pseudorapidity*)

$$\eta \equiv -\log\tan\frac{\theta}{2}$$

The choice of the (r,η,ϕ) coordinate is particularly convenient. Indeed, physics processes are invariant under rotation around the unpolarized beam, ϕ is a natural choice. On the other hand, although the p and $\bar{p}$ have the same momentum in the lab reference frame, their partons carry only a fraction of it. The possible unbalance along the two interactive parton longitudinal components of the momentum, leads to an unknown boost along the z axis of the observed physics interaction. The quantity (called *rapidity*)

$$Y \equiv \frac{1}{2}\log\frac{E+p_z}{E-p_z}$$

is invariant under a boost β along the z axis and the pseudorapidity turns out to be its relativistic approximation ($E \sim p, E \gg m$) making η a useful choice as a detector coordinate in a $p\bar{p}$ collider.
Reflecting the fact that the distribution of the particles generated in the collisions is flat in the (Y-ϕ) space (for the aforementioned symmetry), detector components - whenever appropriate - are uniformly segmented in η and ϕ.
A detailed description of the CDF [1] detector can be found in [14]. In the fol-

[1] We omit on the rest of this document the "II" part of the CDFII (SVXII,...). We always

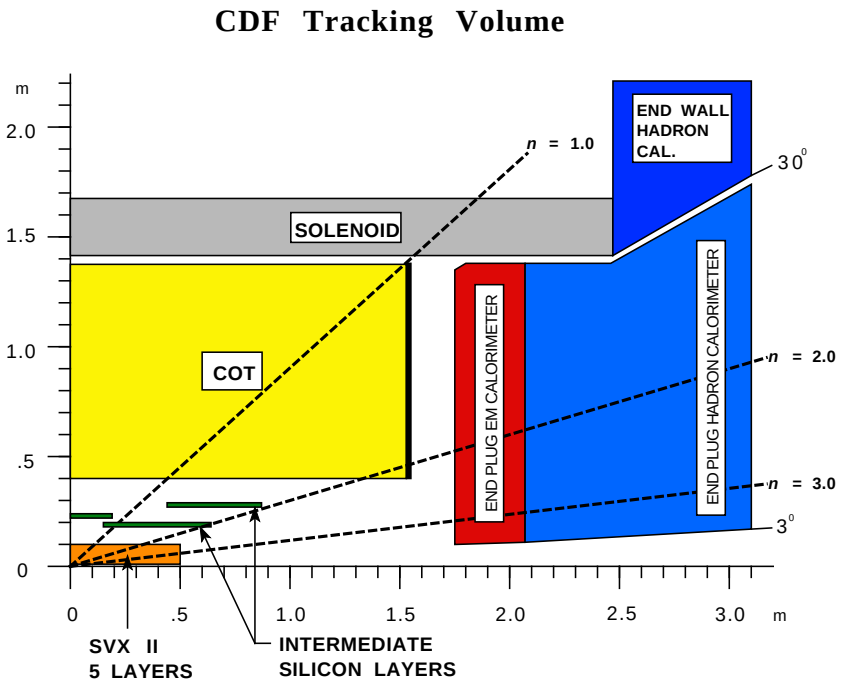


Figure 3.3: *Longitudinal view of the CDF Tracking System for Run II*

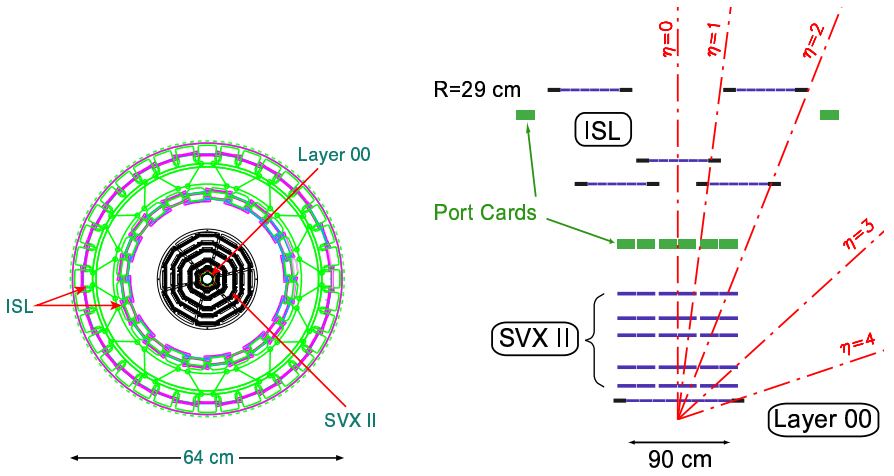
lowing sub-sections we give an overview of the different detector components and their performance, with an emphasis on the elements relevant to the analysis described in this dissertation. A short description of the Trigger and Data Acquisition system is also provided.

3.2 Tracking System

The CDF tracking system is a powerful magnetic spectrometer which consists of several components immersed in a 1.4 T solenoidal magnetic field, parallel to the beam axis. The solenoid coil is built from Al-stabilized NbTi conductor and is located between the Central Tracking System (COT) and the electromagnetic calorimeter (fig 3.1). It operates at the temperature of about 4.7K with a 5KA current, providing a uniform (within 0.1%) field strengh in the region $|z| < 150$ cm and $|r| < 150$ cm.

There are two primary tracking detector systems (fig 3.3): the inner tracking system of the silicon vertex detectors and the outer tracking system of a wire-based drift chamber.

refer to the detector components actually operating in Run II, whether they are the same used in Run I or upgraded or completely new parts.



(a) $r-\phi$ view (b) $r-z$ view

Figure 3.4: *The inner tracking CDF II silicon system*

The later covers the central region in the range $|\eta| \lesssim 1$ with high tracking resolution and reconstruction efficiency. The silicon tracker can instead precisely reconstruct the track impact parameter and as well perform a *"stand alone"* track reconstruction for the region not covered by the COT ($1.0 \lesssim |\eta| \lesssim 2.0$)

3.2.1 Inner Tracker: the Silicon Detectors

The silicon tracking system is part of the CDF detector closest to the interaction point. It consists of three concentric apparatus: the "Layer 00" (**L00**), the Silicon Vertex Detector (**SVX**) and the Intermediate Silicon Layer (**ISL**). The $r-z$ and the $r-\phi$ view of those system are shown in fig 3.4.

This new silicon system was designed to measure with excellent spatial precision tracks in the vicinity of the beam line and at the same time to be able to sustain the large dose of radiation resulting from a long run period. An improved pattern recognition is also implemented along with a faster readout scheme.

Usually a signal is detected on a small cluster of strips and the position is extracted by a center of gravity technique. With this method the accuracy of the individual hit position measurement achieved at CDF is about 12 μm.

SVX The CDF Silicon VerteX Detector, SVX, is built in three cylindrical barrels each 29 cm long. Each barrel supports five double-sided layers between radii of 2.5 and 10.7 cm: three with an $r-\phi$ measurement on one side with the 90° stereo ($r-z$) on the other, and two with an $r-\phi$ measurement with small angle stereo at 1.2°. This design permits a good resolution in locating the z-position of the secondary vertices and to enhance the 3-D pattern recognition

capability of the silicon tracker. Overall there are 405504 channels in the system and the electronics needed to be attached physically on the system. The main drawback of this configuration is the significant amount of heat ($\gtrsim$1KW) that required to incorporate cooling channels into the bulkhead.
It is extremely crucial to be sure to have the SVX barrel placed in a coaxial manner, and it is more important to have the SVX symmetrically around the beam than the z axis of the detector, in order to avoid ϕ dependent measurement of the impact parameter d_0 [2]. The resulting offset between the SVX and COT [3] axis is taken into account using an alignment table that allows the switch between the two references.
Presently the resolution on d_0 is about 40μm, including a 25/30 μm contribution from the beam width. The position on the z axis (z_0) has instead a remarkable resolution of 70 μm.
L00 The L00 detector was added in Run II to improve the impact parameter resolution for high momentum tracks to 25μm. It is mounted just outside the Berillium beam pipe at 1.35-1.62 cm radius and it has an electronics identical to those used in SVX. Its strips provide complete ϕ coverage, and z coverage extending $\pm$78.4 mm from $z = 0$. It is composed of a single-side silicon p-in-n silicon wafer allowing a greater radiation resistance from the intense environment close to the interaction point.
ISL Besides the L00, another silicon detector was added in Run II, the Intermediate Silicon Layer (ISL). It consists of 5 separated layers situated at radii of 20, 22 and 28 with the configuration showed in fig 3.4 with one layer in the central region and two for $1< |\eta| <2$. Each layer is made up of double-side microstrip as in the SVX with a stereo angle. The double-side microstrips, with a 1.2° stereo angle, are read-out only every other strip to reduce the total channel count to 286,800. Thus, the single hit resolutions is 16 (23) μm on the axial (stereo) side, compared to the 12 μm for the SVX. The ISL improves tracking in the central region and in conjunction with the SVX allows *stand-alone* silicon tracking in the region $1< |\eta| <2$, expanding the Run I tracking coverage.

3.2.2 Central Outer Tracker (COT)

The Central Outer Tracker (COT) [15] is a 96-layers open cell drift chamber used for charged particle tracking in the $|\eta| < 1.1$ region. The COT surrounds the silicon detectors and extends to a radius just inside the Time-of-Flight system (fig 3.1). The design goal of the COT is to reproduce the functionality of the CTC, which was used in Run 0 and Run I, but with a faster drift time (less than 100 ns), using smaller drift cells and a fast gas. The gas mixture is Ar-Ethan-CF_4 50:35:15 with a drift velocity of $\sim$ 100 μm/ns. The COT 96 sense wire layers are radially grouped into eight "superlayers" (fig 3.5), each one divided

[2] d_0 is particularly critical because is used to filter the data (trigger)
[3] It is important to mention that the global coordinate system described in sec 3.1 has the z axis coincident with the z COT cylinder axis, *bydefinition*. The local coordinates of the other components of the detector are converted in CDF global coordinate using the measured relative position with respect to the COT.

in into Φ "supercells" (fig 3.6), and each supercell has 12 sense wires and a maximum drift distance that is approximately the same for all superlayers.

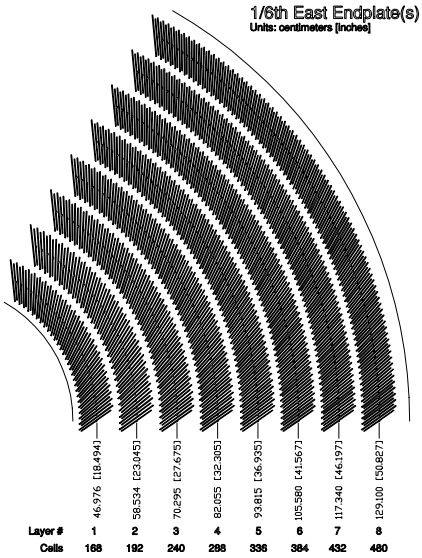


Figure 3.5: *One eight of the east COT end plate showing the wire-plane slots grouped into eight superlayers.*

The entire COT contains 30,240 sense wires. Approximately half the wires run along the z direction ("axial"). The other half are strung at a small angle ($2°$) with respect to the z direction ("stereo"). The active volume of the COT is located at a radius between 43 and 133 cm from the nominal beamline and the chamber is 310 cm long. Charged particles passing through ionize the gas, and the electrons produced drift toward the sense wires because of the electric field generated by the potential wires. Secondary ionizations generated by the $1/r$ sense wire electric field produce a signal (hit) further amplified and eventually recorded by the electronics attached at the end of the wire. The $r-\phi$ position of the track with respect to the sense wire is then inferred from the arrival time of the signal.
Because of the presence of a magnetic field, the electrons move at an angle α (Lorentz angle) with respect to the electric field lines. The value of α, depending on the magnitude of both fields and the properties of the gas mixture, is $\sim 35°$. Therefore, to maximize the resolution for the almost radial high p_T tracks, all cells are tilted by $35°$ (fig 3.5). Moreover this tilt allows wire planes overlap in radial view, which means more hits for high p_T tracks. Finally the large tilt helps in resolving the left/right ambiguity of the position of the hit. Track reconstruction is made matching track segments starting from the axial superlayers and reconstructing the $r-\phi$ projection. Then the result is compared with the $(r-\phi)$ plane hits from stereo superlayer and from the resulting offset/mismatch

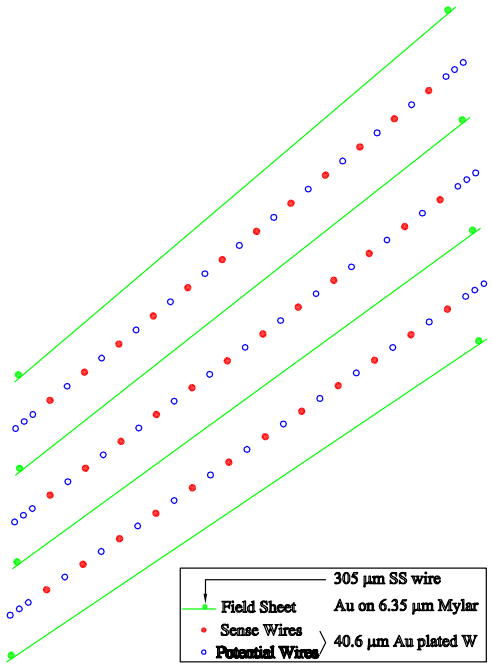


Figure 3.6: *Close-up view of three cells from super layer 2 of the COT which illustrates the position of the wires with respect to the field sheets.*

the longitudinal parameters of the helix (z_0 and $\cot\theta$) are measured.
The single hit position resolution has been measured to be [4] ~ 140 μm which can be translated into a transverse momentum resolution of [5]

$$\frac{\delta p_T}{p_T} \sim 0.15\% \frac{p_T}{[GeV/c]}$$

After a track is reconstructed in the COT (called *"COT-only"* track), the information from the SVX is added to achieve the most complete helix information. The COT-only track (that comes with an error) is extrapolated to the outermost layer of the SVX and each hit within a certain radius is appended to the track and a re-fit is performed. A new candidate track is generated for each appended hit, but only the best two, in terms of fit quality and number of hits, are kept. Both tracks are then extrapolated further in, where the process is repeated. At the end the best among all the track candidates associated to the original COT-only track is promoted to be the "default" track.

[4] This is the most up-to-date number for the hit resolution obtained with the latest tracker alignment corrections in the CDF offline software

[5] The p_T resolution for Run I CTC chamber was $\sim 0.2 \frac{p_T}{GeV/c}$

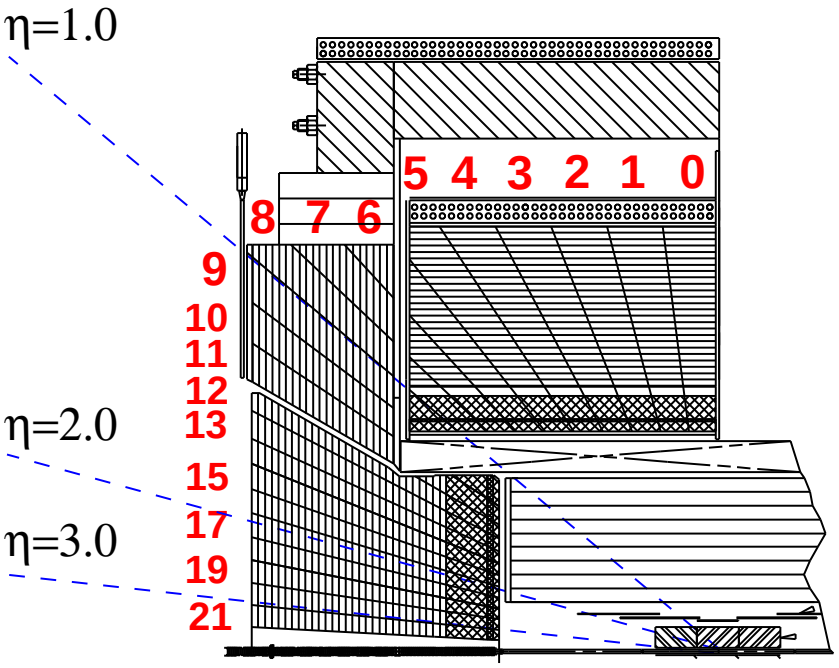


Figure 3.7: *A quarter view of the CDF calorimeter systems and its segmentation. Also the tower number is reported. It is important to notice that although most of the towers have about the same size in* η *(*$\sim$ *0.1), those ones beyond* $|\eta| \gtrsim 2$ *are considerably larger (for the last tower* $\Delta\eta \sim 0.6$*)*

3.3 Calorimeter Systems

Surrounding the tracking volume and solenoid, the CDF calorimeter modules measure the energy of particles produced by the $p\bar{p}$ collisions.
In Run II, the gas calorimeters in the forward region ($|\eta| > 1.1$) used in Run I were all replaced with a new scintillating tile plug calorimeter. The CDF calorimeters are now all scintillator-based sampling calorimeters - based upon sandwiching scintillating material between layers of heavy material - and have a uniform pattern of matched projective towers of Electromagnetic (EM) and hadronic (HAD) compartments (as shown in fig 3.7).

When a particle travels through the calorimeter, a particle shower made of electrons, photons and hadron is produced - at the expenses of the particle energy - as a consequence of its interaction with the material. When the shower passes through the scintillator, photons emitted by atomic excitation are collected by photomultiplier tubes (PMT) and converted in electric charge. Eventually all the original particle energy is released inside the calorimeter and measured integrating the charge collected in the PMTs.
The calorimeter system is divided into a central (covering the region with

$|\eta| < 1.1$) and a plug calorimeter. A summary of the sub-system properties and coverage is given in Table 3.2 while the details are discussed in the following paragraphs.

	CEM	**CHA**	**WHA**	**PEM**	**PHA**
$\|\eta\|$ coverage	< 1.1	< 0.9	$0.7< \|\eta\| <1.3$	$1.3< \|\eta\| <3.6$	$1.3< \|\eta\| <3.6$
Modules	48	48	48	24	24
η towers per module	10	8	6	12	10
Total Channels	956	768	676	960	864
Absorber(inches)	pb (0.13)	Fe (1)	Fe (2)	pb (0.18)	Fe (2)
Thickness	$19X_0$, $1\lambda_I$	$4.5\lambda_I$	$4.5\lambda_I$	$21X_0$, $1\lambda_I$	$7\lambda_I$
Position Res.	0.2×0.2	10×5	10×5		N/A
Energy Res.	$\frac{13.5\%}{\sqrt{E_T}} \oplus 1.7\%$	$\frac{75\%}{\sqrt{E_T}} \oplus 3\%$	$\frac{80\%}{\sqrt{E_T}}$	$\frac{16\%}{\sqrt{E_T}} \oplus 1\%$	$\frac{80\%}{\sqrt{E_T}} \oplus 5\%$

Table 3.2: *Design parameters and characteristics of CDF calorimeters. The position resolution (in ($(r-\phi)\times z$ cm) is measured at 50 GeV incident particle. CEM resolutions obtained using CES chambers.*

3.3.1 Central Calorimeter

The central calorimeter is the same detector used by CDF in Run I. It is a cylindrical symmetric detector divided in half at $\eta = 0$. The two halves are then segmented into 24 wedges of 15° in ϕ for a total of 48 modules. Finally, each wedge is divided into 10 projective towers such that each tower has $\Delta\eta \sim 0.1$ (fig 3.8).
Different kind of particles interact in different way when they pass through the matter. Typically the energy loss is parametrized with a characteristic length. For the electromagnetic particles (photons and electrons) it is the radiation length X_0 and for hadronic particles the interaction length [6] λ_I. Each wedge is composed by an electromagnetic compartment (CEM) made of high-Z absorber - X_0 grows with Z - to detect photons and electrons, and a larger outer hadronic compartment made of relatively inexpensive low-Z absorber to detect hadronic particles.
It contains 31 layers of 0.125 in lead interleaved with 5.0 mm polystyrene scintillator giving a total radiation length of 18 X_0, enough to stop most of the EM particles produced in the collision. Some very energetic photon or electrons though can punch through the electromagnetic compartment and reach the hadronic calorimeter. The energy resolution for the CEM is measured to be $\sigma/E = 13.5\%/\sqrt{E}$ [16].

[6] In particular electrons lose their energy via bremsstrahlung, photons via pair production and hadrons by nuclear interaction

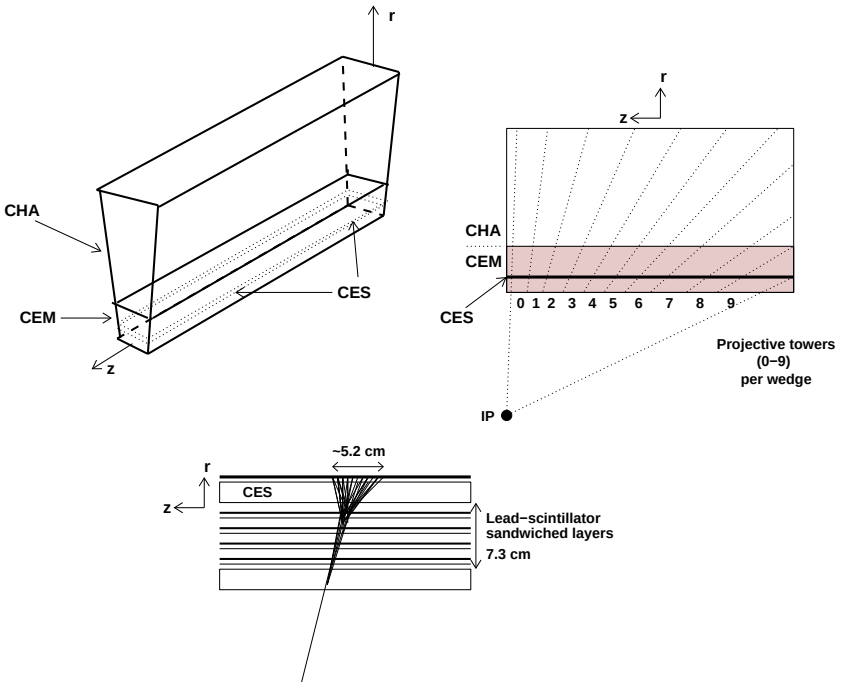


Figure 3.8: *Top: A view of one central calorimeter tower wedge with its segmentation and tower number. Bottom: CES position in the electromagnetic calorimeter and its response to an EM incident particle.*

The central and end wall hadronic calorimeter (CHA, WHA) use iron as radiator and PMMA napthtalane scintillator [17]. The CHA (WHA) has a structure of 32 (15) layers and each layer is composed of 2.5 (5.1) cm of iron absorber and 1.0 (1.0) cm of plastic scintillator. The η ranges are reported in tab 3.2. The total hadronic absorbing length is $\sim 4.7\lambda_0$ for both CHA and WHA with the energy resolution for the CHA of $\sigma/E = 50\%/\sqrt{E} \oplus 3\%$ whereas for the WHA is a little worse, $\sigma/E = 75.0\%/\sqrt{E} \oplus 4\%$

The Shower Max Detector

At the most probable location of the shower maximum (after about 6 radiation lengths) a Ar/CO_2 proportional strip detector (CES) [16] has been embedded in each CEM wedge (fig 3.8). It determines with high precision the shower position and its transverse development by measuring the charge deposition on orthogonal strips and wires. Cathode strips running in the azimuthal direction provide z information, while anode wires running in the z direction provide $r-\phi$ information.

The CES is particularly precious to measure, with good accuracy, the isolated photons position that, being neutral, cannot be determined with the tracking system[7].
The transverse shower shape information instead is used to distinguish a single particle shower from a multi-particle final state. This plays an important role in discriminating the multiple photon decays of neutral mesons - like π^0 and η^0 - from prompt photon final state (sec 4.8.1). Recently also it has been shown [45] that CES information can be successfully used to detect the early shower in the EM calorimeter of hadronic particles, an information that is of great help in separating interacting hadrons and minimum ionizing particles in the CEM. The basic orientation of the strips and wires are shown in fig. 3.9. The CES itself is located 184 cm from the beam line, and has a total thickness of 0.75". There are two wire sections per wedge - corresponding to towers (0:4) and (5:9) - with 69 ($0< |z| <121.2$ cm) + 59 ($121.2< |z| <239.6$ cm) strips and 64 wires readout channels[8]. The position resolution for the CES is approximately 2 mm in each direction for a 50 GeV electron.

The Central Preshower Detector

In front of the CEM wedge another multiwire proportional chamber, the Central PReradiator Chamber (CPR), samples the early development of the electromagnetic showers that begins in the solenoid magnet material [9]. This is exploited to enhance the purity of electron and photon selection suppressing the hadronic background. The CPR is positioned at a radius of 168 cm from the beamline

[7]Of course the CES works as well with electrons and track/CES matching is actually used to suppress fakes

[8]There are actually 62 wire cells each half-section, but all the wires - except the ones the edges - are ganged in pairs resulting in a 64 channel readout

[9]The thickness of the solenoid is $1.075X_0$

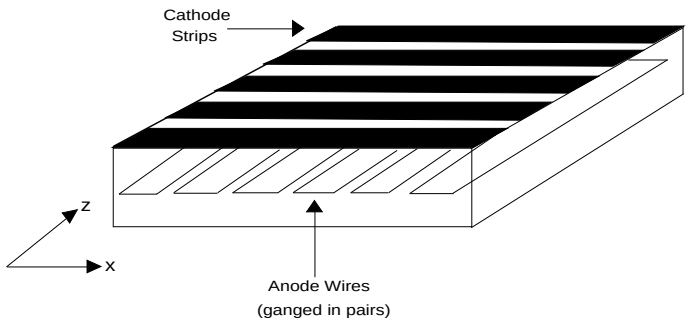


Figure 3.9: *A sketch of the Central Shower Max Detector. The cathode strips run in the x direction providing z information while the anode wires run in the z direction providing x (i.e. ϕ) information.*

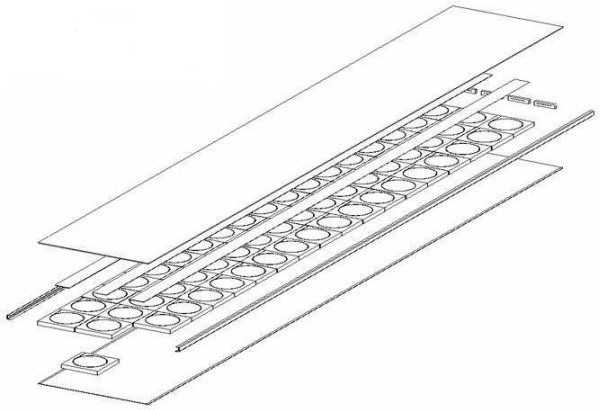


Figure 3.10: *Schematic drawing of one CPR II module assembled in aluminum shell (a.k.a. "pizza-pan"). The system is composed by a* 3×18 *tiles array (*12.5×12.5 cm^2 *each) and a WLS fiber is embedded into a groove carved on the surface of each tile*

between the CEM and the solenoid (see fig 3.1). The chambers have 2.22 cm cells, segmented in $r - \phi$, providing $r - \phi$ view from the wire readout. There are 4 chamber divisions spanning ± 1 in pseudorapidity. They have 33 ground wires alternating with 32 sense wires. Overall there are 16 readout channels per chamber and 32 channels per wedge.

The New Preshower Detector

The wire aging of the CPR gas detector, the small signal pulses and the poor response resolution suggested to replace the operating CPR with a better detector [19]. This happened during the fall 2004 Tevatron shutdown, when a new preshower detector based on scintillator tiles and read out by wavelength-shifting (WLS) fibers have been installed in the same location of the old CPR. The new CPR consists of 48 modules, each one covering the front face of one calorimeter wedge and made of 54 20 mm thick scintillator tiles (see fig 3.10).

A 1mm diameter WLS fiber embedded in a groove curved on the surface of each tile serves as read-out. This configuration assures a much finer segmentation with respect to the old CPR needed to suppress the enhanced underlying energy contribution brought about the Run II high luminosity regime [10].
At the same time a very similar but smaller detector (48 modules, 10 tiles per module) was installed between the wedge to replace the old, seldom used [11] Crack Chamber Detector (CCR). The new CCR is basically integrated with the

[10] Indeed the old gas-based CPR was expected to be useless for $\mathcal{L}_{inst} > 2 \cdot 10^{32}$ because of the overwhelming underling events

[11] The old CCR was composed by a 12 X_0 tungsten bar paired with a wire chamber. Unfortunately the resolution and noise of the system was too poor to be useful in the energy measurement although they have been valuable in studying the occasional odd event.

CPR detector with identical electronics and read out structure. The main purpose of this detector is to recover part of the energy lost in the not instrumented regions between the calorimeter wedges (the ϕ-cracks, each about 1/15 of wedge wide) allowing a more precise evaluation of the missing energy.

3.3.2 Plug Calorimeter

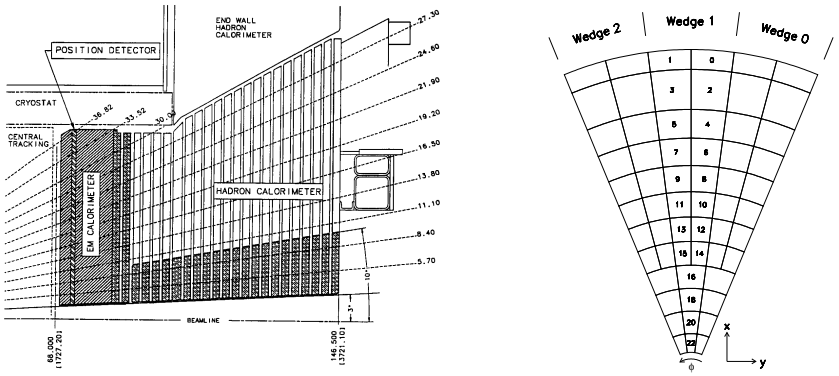
The two old gas End Plug and Forward Calorimeter used in Run I have been replaced with a new lead/scintillator sampling calorimeter, a much faster detector suitable to deal with the higher Run II Tevatron event rate.
The Plug (or forward) calorimeter has two sectors [18], the electromagnetic (PEM) and the hadronic (PHA) component and it matches the segmentation and the projectivity of the central calorimeters extending the coverage until $|\eta| \sim 3.6$ (fig 3.7). The segmentation in ϕ is finer than in the central with towers covering 7.5° for low η region and 15° for high η (fig 3.11). The electromagnetic section is constructed from 4.5 mm lead and 4.0 mm scintillator, with 23 layers providing a total radiation depth of 21 X_0. The hadronic section also has 23 layers but is made of 2.0 in iron and 6.0 mm scintillator with an absorption length of 7.0 Λ_0. The design gives an EM energy resolution of $\sigma/E = 16\%/\sqrt{E} \oplus 1\%$ for single electrons, and an hadronic resolution $\sigma/E = 80\%/\sqrt{E} \oplus 5\%$ for single pions (see Table 3.2 for a summary of the Plug specifics).
As in the central calorimeter, a shower-maximum detector (PES) is located about 6 X_0 deep within the PEM, and it consists of two layers (denoted 'U' and 'V') of scintillating strips. The strips are 5 mm wide, and roughly a square in cross section. The PES is segmented into 45° sectors, with strips in the U and V layers offset from the radial direction by +22.5° and -22.5° respectively. Position resolution of the PES is 1 mm. There is also segmentation in pseudorapidity, with a low-η region (1.13< $|\eta|$ <2.60) and a high-η region (2.60< $|\eta|$ <3.50).
There is not a separated preshower detector in the Plug region, but the first layer of the scintillator is made thicker to yield more light and is read out separately. By analogy with the CPR this system is called PPR.

3.4 Other systems

In this section we review the most significant of the remaining systems. Because these are not directly involved in the analysis presented, only a brief summary is given.

3.4.1 Time of Flight

A Time-of-Flight (TOF) [20] detector has been added to CDF for Run II in order to expand the CDF particle identification capability in the low p_T region. The TOF system is indeed designed to distinguish low momentum pions, kaons and photons by measuring the time it takes these particles to travel from the primary vertex of the $p\bar{p}$ collision to the TOF system (fig 3.12).



(a) Plug Section (b) 45° Plug Slide

Figure 3.11: *The Plug calorimeter schematic (a) with a view of its tower geometry (b). As shown, the region 1.1< $|\eta|$ <2.11 has a $\Delta\phi = 7.5°$ whereas the region 2.11< $|\eta|$ <3.6 has a $\Delta\phi = 15°$. The wedge definition is less important for the Plug because - in contrast with the central calorimeter - there is not gap (ϕ-crack) between them.*

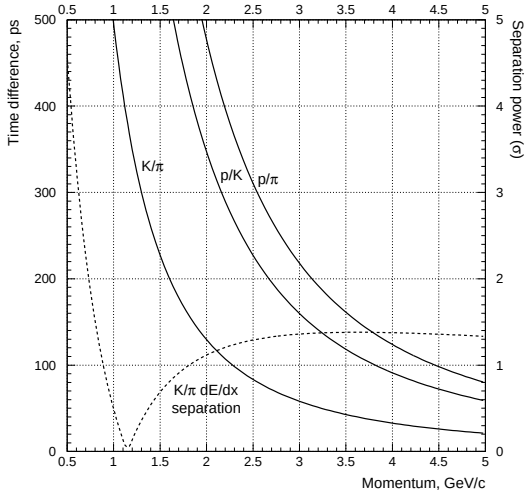


Figure 3.12: *Time difference as a function of momentum between $K/\pi, p/K$ and p/π traversing a distance of 140 cm, expressed in ps and separation power, assuming a resolution of 100 ps. The dashed line shows the K/π separation power from the dE/dx measurement in the COT.*

It consists of 216 scintillator bars installed at a radius of about 138 cm (from the z axis) in the 4.7 cm space between the outer shell of the COT and the

cryostat of the superconducting solenoid (see fig 3.1). Bars are approximately 279 cm long and 4x4 cm^2 in cross section providing 2π coverage in ϕ and roughly (-1;1) coverage in η. Bars are read out at both ends by fine-mesh PMTs capable of providing adequate gain even if used inside 1.4 T magnetic field. Usage of long attenuation length fast rise time scintillator in conjunction with these PMTs allowed to achieve specified resolution of 100 ps. More details on TOF and its performance can be found in Reference [20]

3.4.2 Muon Detection System

Muons are particles interacting with matter only by ionization but, because of their high mass, they do not cause showers in the calorimeters [12]. Therefore a muon detection system is installed radially outside the calorimeter. In order to suppress the background coming from secondary interactions in the far forward parts of the beam line and from residual hadrons surviving the passage in the calorimeter, absorbers are located - whenever possible - in front of the muons detectors [13] (calorimeter steel, the magnetic return joke, additional steel walls and the steel from the Run I forward muon toroids [14]).
A muon candidate is detected for a track segment it leaves in the muon chambers (muon *stubs*). Timing information is used to reject background and a matching between the stub and the extrapolated track reconstructed in the COT is performed. A χ^2 value is computed for the track-stub match and eventually used to define good quality muons.
There are four muon subsystem at CDF consisting of sets of drift chambers and scintillators: the Central MUon Detector (CMU), Central Muon UPgrade Detector (CMP), Central Muon EXtension Detector (CMX) and Intermediate MUon Detector (IMU). The central part of the system is similar to Run I, except upgrades to improve readout and geometrical coverage. The IMU is new to Run II. The $\eta - \phi$ coverage of the muon system for Run II is shown in fig 3.13 where is compared with the Run I coverage. The design parameters are outlined in Table 3.3.

3.4.3 Cherenkov Luminosity Counter

The Cherenkov Luminosity Counter [21], CLC, measures the average number of interactions per bunch crossing μ. Then instantaneous luminosity $\mathcal{L}_{inst}$ is extracted using the formula

$$\mu \cdot f_{bunch} = \sigma_{p\bar{p}} \cdot \mathcal{L}_{inst} \tag{3.1}$$

[12] At least for energy relevant to this experiment

[13] Reducing this background is particularly important to keep under control the rate of the muon triggers where the information from these detectors are actually used

[14] Although heavy shielding improves the muon detection purity, it causes the losing of their energy as well. Therefore the minimum p_T threshold for muon detection depends upon the amount of material in front of the particular muon system

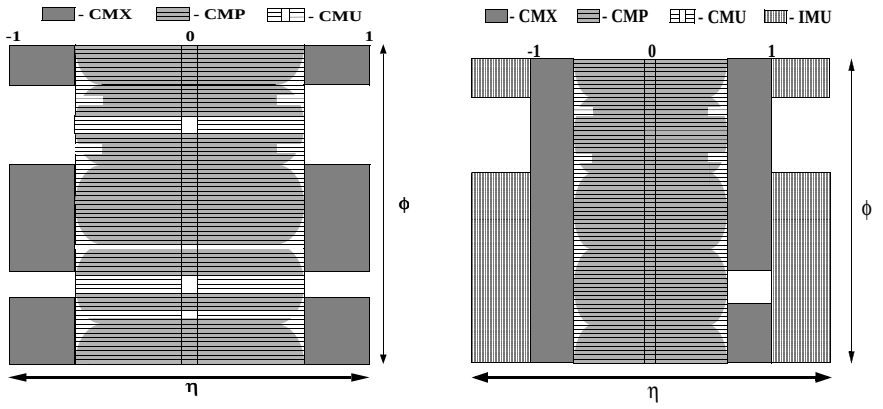
where $\sigma_{p\bar{p}}$ is relatively well known total $p\bar{p}$ cross-section at $\sqrt{s}$ =1.96 TeV and f_{bunch} is the rate of bunch crossing in the Tevatron. There are two CLC detector modules installed in the pseudorapidity range $3.7 \lesssim |\eta| \lesssim 4.7$ (see fig 3.14). Each CLC module consists of 48 thin, long, conical, gas-filled, Cherenkov counters. They are arranged around the beampipe in three concentric layers with 16 counters each, and pointing to the interaction region. The two outer layers are 180 cm long and the inner layer 110 cm whereas the diameters range between 2 and 6 cm.
A primary particle from $p\bar{p}$ interaction will transverse the full length of the counter generating a large amount of light collected by the PMTs located at the far end of each counter. On the other hand secondary particles produced on the beampipe and materials surrounding the CLC, will cross the counters at different angles yielding much smaller signals.

Thus, being the PMT signal collected by the Cherenkov counter proportional to the primary particle number, it is possible with the CLC to discriminate multiple interactions within the same bunch crossing and measure with high accuracy μ even in an high luminosity regime [15]. The final luminosity measure-

[15]Traditional scintillator-counter-based luminosity measurement uses the fraction of empty

	CMU	**CMP/CSP**	**CMX/CSX**	**IMU**
η coverage	$\lvert\eta\rvert \lesssim 0.6$	$\lvert\eta\rvert \lesssim 0.6$	$0.6 \lesssim \lvert\eta\rvert \lesssim 1.0$	$1.0 \lesssim \lvert\eta\rvert \lesssim 1.5$
Drif Tubes:				
thickness (cm)	2.68	2.5	2.5	2.5
cross section (cm×cm)	2.68 x 6.35	2.5 x 15	2.5 x 15	2.5 x 8.4
Max drift time (μs)	0.8	1.4	1.4	0.8
Total drift tubes (Run I)	2304	864	1536	none
Total drift tubes (Run II)	2304	1076	2208	1728
Scintillator:				
thickness (cm)	-	2.5	1.5	1.5
Area (width×length) (cm×cm)	-	30-40 x 180	17 x 180	30 x 320
Total Counter (Run I)	-	128	256	none
Total Counter (Run II)	-	274	324	864
Pion Interaction length	$5.5\lambda_I$	$7.8\lambda_I$	$6.2\lambda_I$	6.2 - 20 λ_I
Minimum Muon P_T (GeV/c)	1.4	2.2	1.4	1.4 - 2.2
Multiple scattering res. (cm/p_T)	12	15	13	13 - 25

Table 3.3: *Design parameter for the CDF muon detectors. Pion interaction lengths and multiple scattering are quoted for a reference angle of $\theta = 90^\circ$ in CMU nd CMP/CSP, for an angle of $\theta = 55^\circ$ in CMX/CSX and show the range of values for IMU.*



(a) Run I (b) Run II

Figure 3.13: *Schematic drawing of the Muon System coverage in Run I (a) and in Run II (b).*

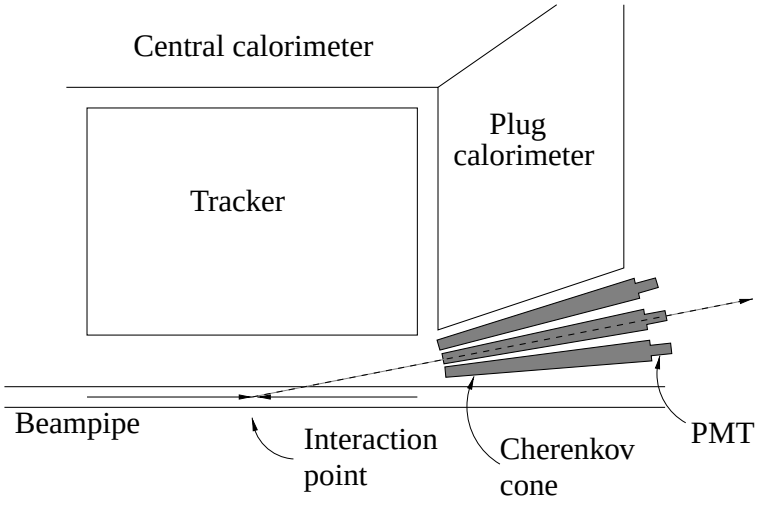


Figure 3.14: *A schematic view of the CLC and its location in the CDF detector.*

ment is eventually evaluated correcting offline the CLC outcome analyzing the Minimum and Zero Bias trigger data.

crossing to measure the number of interactions in a bunch crossing. However at high luminosity the no-crossing fraction became too small to achieve a relative small uncertainty.

3.4.4 Forward Detectors

The CDF Forward Detectors were designed to enhance the sensitivity for the Run II very forward physics program. They include the Roman Pot Fiber Track Spectrometer (RPS), a set of Beam Shower Counters (BSCs) and two forward MiniPlug Calorimeters (MP). They are literally an extension of the CDF detector since the BSCs - except two of them - and the RPS are not located in the CDF collision hall but inside the Tevatron tunnel (see fig 3.15).

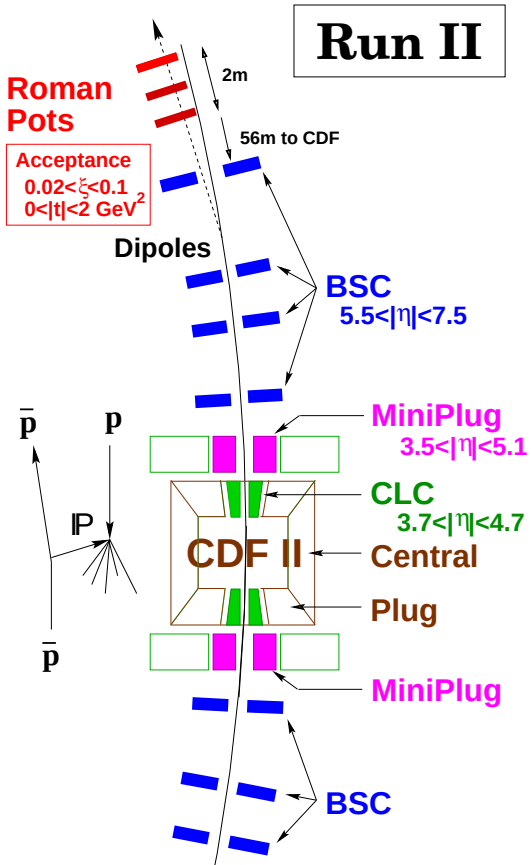


Figure 3.15: *Forward Detector location along the beam-pipe (not in scale)*

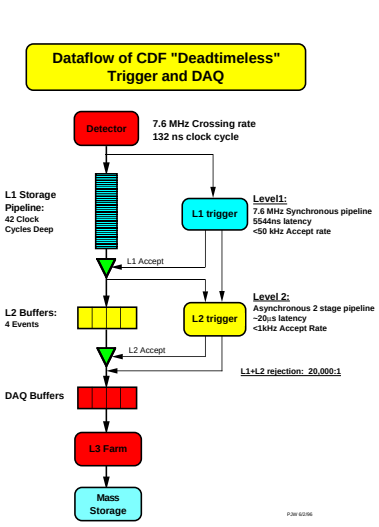
The RPS were installed for the Run IC data taking and their purpose are to measure the momentum of the leading antiproton. It is a fiber detector spectrometer consisting of three station, 1 m apart from each other. The fiber detector reads X and Y (40 channels each) coordinates to identify the position of the tracks with a resolution of approximately 100 μm from which the $\bar{p}$ momentum is determined. No change to the detector has been made for Run II, but the readout electronics was completely redesigned to take into account the shorter Run II bunch crossing time.

The BSCs are scintillator counters capable to detect particles traveling in either direction from the interaction point along and near the beampipe and cover the pseudorapidity region $5.5 < |\eta| < 7.5$. There are four BSC stations on the west side and three on the east side of the interaction point (fig 3.15). Each station are made of two scintillation counters, except for the BSC-1 stations, which have four counters. They are mainly used as rapidity gap tagger - i.e. in a "veto" mode - rejecting non-diffractive events when a signal above threshold is detected. The BSC-1 counters are also used to monitor Tevatron beam losses and collision rates providing a beam quality index to be used for selecting appropriate beam condition.
The MiniPlug [22] detectors are two small calorimeters (placed one on East and one on the West part of the detector) covering the region $3.5 < |\eta| < 5.1$. They are very different from the other CDF calorimeters both in size, in design and in resolution built with the main purpose to measure the event energy flow in the very forward direction. They are composed by only one compartment to detect both electromagnetic and hadronic calorimeter showers. They consist of alternating layers of lead plates and liquid scintillator read out by wavelength shifting (WLS) fibers. The WLS fibers are perpendicular to the lead plates - parallel to the proton/antiproton beams - and uniformly distributed, generating thus a "towerless" geometry and no dead regions due to the lack of internal mechanical boundaries. For read out purpose the fibers are conceptually grouped in small subsets called, again, towers. The active depth of each MP is 32 radiation lengths and 1.3 interaction lengths. The energy resolution, as measured with a small-scale prototype in a test beam study, is $\sigma/E = 18\%/\sqrt{E}$ for single electrons.

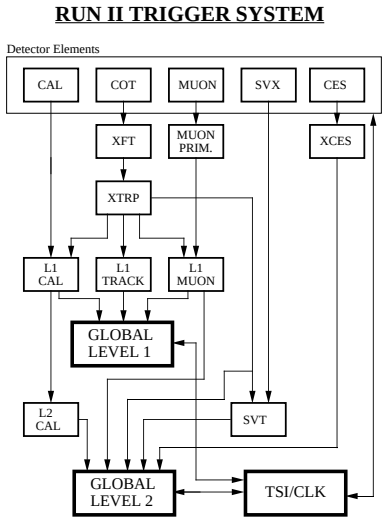
3.5 Trigger and Data Acquisition System

The 36 bunches of protons and antiprotons traveling along the Tevatron ring cross every 396 ns, corresponding to 2.5 million events every second (2.5 MHz). At the present instantaneous luminosity there is about one interaction per crossing. This huge number of events is by far beyond the CDF mass storage capability. Indeed, each event requires about 250 KB to be stored and it would be necessary a bandwidth of 630 GB/s that is not only very difficult and expansive to handle, but it would also require a stunning amount of disk space impossible to have. Moreover it takes about 2 ms to read the entire detector out, when in the meantime more than 5,000 other events would go unrecorded. To overreach these problems and to not exceed the current data acquisition limitation, an online system selection is implemented in the Data Acquisition System (DAQ) [23] called *trigger*. It acts as a filter on-the-fly, disregarding the majority of the events that not meet basic set of requirements to be considered interesting for physics analysis while sending the others out to be stored. Because of the high beam crossing rate, it is not possible as in Run I (when the bunch crossing separation was 3.5μs) to make a decision whether to keep the events or not in the 396 ns between two bunches. Indeed a fully pipelined data acquisition system

and a three level trigger was designed (fig 3.16(a)). The sophistication of the filtering process and the time required to accomplish increase at each stage of the trigger and eventually only 30-50 events of the original 2.5 million are stored on tape every second. The structure of the three levels of the trigger systems follows.



(a) CDF readout functional block diagram



(b) CDF trigger system block diagram

Figure 3.16: *Block diagram detailing CDF DAQ and L1/L2 trigger system*

Level 1

The first decision and filtering level of CDF is the Level 1 trigger. It has been designed to reduce the rate from 2.53 MHz to 50 kHz. To give the trigger the time to make its decision, all front-end electronics are equipped with buffers 42 clock cycles deep [16]. Every clock cycle the event is moved up one slot in the pipeline until either a decision is made (accept/reject) or it reaches the end of the pipeline. In this case the event is lost. Therefore the Level 1 has about 5,554 ns to make its decision. It uses rudimentary (not detailed reconstructed) objects called *primitives*. Information from the calorimeter, COT and the muon detectors fed three possible streams (fig 3.16(b)). The calorimeter stream (L1CAL)

[16] All the Run II electronics was designed to handle 192 ns bunch crossing run condition as preparation for the Tevatron upgrade scheduled at same point during the Run II (the so called Run IIB). Later, this upgrade has been canceled. Thus to collect the data at every 396 ns, every other cycle is automatically rejected. This means that the L1 storage pipeline is basically 14 (396 ns) bunch crossing deep.

bases its decision upon the energy deposited in the calorimeter towers [17] to identify objects (electrons, photons and jets) or global event variables (missing energy and total energy). The eXtremely Fast Track (XFT) performs a rough reconstruction of the COT tracks and takes a decision based on the number and transverse momentum of those tracks. Moreover, the extrapolated track information (XTRP) is sent to the calorimeter and muon stream to identify electrons and muons.
The muon stream uses the information from the muon chamber and scintillator systems along with the XRTP to trigger on muons.

Level 2

The events passing the L1 trigger proceed to the Level 2. They are written in one of the four data buffers - present in each detector component electronics - waiting to be scrutinized. The L2 is an asynchronous system which processes events in a time-ordered fashion. As opposite to the L1 pipeline, data remain in the L2 buffer until a decision is made.If all the four spots in the buffer are occupied, no more events from L1 can be accepted and they are lost ("dead-time"). To minimize the deadtime, the latency of the L2 decision is 20μs and it uses the L1 primitives as well as additional data coming from the central strip chambers, the silicon vertex and a more sophisticated elaboration of the calorimeter information (fig 3.16(b)).
The CES information (XCES) are used to improve the purity of the electron and photon selection. A SVX 2-D fast tracking is performed (SVT) including a impact parameter d_0 calculation allowing to trigger on secondary vertex in SVX. Also a simple calorimeter clustering (L2 jets) is performed discriminating between hadronic clusters (real jets) and electromagnetic clusters. At a second stage of the L2 trigger decision further event topology characterization and tighter requirements over the aforementioned object are applied through four programmable processors (Alpha). The L2 rejection factor is about 150 and the accept event rate is around 300 Hz.

Level 3

After a Level 2 accept signal is issued, the event information are sent to an Event Builder (EVB), a component of the Level 3 trigger. It collects all the digital information from each detector and arranges event fragments based upon the bunch counter. This step is very important to avoid mixing up information from different bunch crossing. The assembled fragments (*"event record"*) are sent to the L3 processing PCs farm for analysis. It consists of 16 identical sub-farms each composed of 12-16 processor nodes. The system is completely scalable and the processors can be increased/upgraded with minimal impact. Each event record is fed to one of the processor nodes where the event reconstruction is

[17] The segmentation of the calorimeter used in the trigger is different from that one used in the analysis. Typically a "trigger tower" corresponds to two or more "calorimeter" towers

performed and final trigger requirements are applied. The designed latency for the L3 is a full second allowing a whole event reconstruction and the calibration application. Thus the L3 filtering can be performed with almost the same detailed information and algorithms used later in the offline analysis making the L3 trigger process very efficient.
Once an event is accepted, it is sent to permanent storage like tapes ready to be analyzed. The accept rate for L3 is limited to the rate at which data can be written to tape, and it is currently about 75 Hz.

3.6 Data Processing

The Level-3 data flow is segmented into streams written to tape in real time as the data are being taken. After that some other manipulations occur before the data are available to the final user. They are referred as offline data handling. The streams are further sub-categorized in datasets, set of events with common physics properties (usually determined by the L3 criteria). All the events are then reprocessed with the CDF Offline reconstruction code with the operation known as "*production*". This is a very important step because at this stage the raw data are unpacked and (collections of) physics objects suitable for analysis - as tracks, vertices, muon, electrons, photons, jets, etc... - are generated. This is similar to what is being done at Level 3, except it is done in a much more elaborate fashion, applying the most up-to-date calibrations, using the best measured beamlines, the best vertexing algorithm, etc...
It is important to notice that the reconstruction algorithms have been improving as more data area is being analyzed and a better knowledge of the detector performance is being achieved. This is a continuous process and occasionally, when a significant amount of changes have been accumulated, all the data are re-processed again with the new improved *production* code. So one needs to state the production version to fully specify the data used in the analysis. For this analysis we used the `cphjj0d` dataset produced with the 5.3.3 offline version. For consistency all Monte Carlo samples used in this analysis have been generated with the 5.3.3 simulation packages.

Good Run List

The detector data taking is a continuous process that can start once the proton and antiproton beams injected in the Tevatron ring have been stabilized and it ends when the beams are too degraded to produce a reasonable amount of luminosity ("*store*"). At this point the store is dumped and a fresh one is going to be injected to restart the whole procedure again. A continuous period of data taken is called *run*. During the run the detector configuration (i.e. online calibration, trigger table, high voltage setting, etc...) is stable, although it can change from run to another. During the data taking some of the detector components or part of the DAQ may register either a malfunctioned or an anomalous behavior or even a failure. In this case there are two possible actions

[18]: (a) if the malfunction is critical for the all data taking process - it is not possible to take data at all - the run is terminated, and a new run will start after the problem is fixed; (b) if the malfunctioning affects only a detector subsystem but the data taking can continue, the run is not stopped but it is marked *bad* for that particular subsystem [19]. Eventually the whole run can be marked *bad* if the data taking condition requires to do so, otherwise it is marked *good*.
For each analysis a *"Good Run List"* is compiled, based on the good runs having all the detector components, whom information are actually used, marked *good*. In the analysis discussed in this dissertation we selected only the events being marked *good* for the COT, the calorimeters, the CES and the CPR subsystems.

[18]There are actually other possibilities on what can happen during the data taking and what actions undertake that we omitted for brevity. Moreover these protocols change often reflecting the improvements applied to the DAQ software aimed at reducing the data taking downtime.

[19]Sometimes a particular sub-detector is not included in the DAQ system on purpose - even if perfectly working - to avoid possible damages caused by unstable condition of the beam. This happen often with the silicon detectors when the beam losses are too high than a fixed security level.

Chapter 4

Particle Identification

4.1 Event Reconstruction

In order to analyze the physics contents of the debris generated after each proton antiproton collision, the raw information collected by different detector components is assembled together to form meaningful physics objects such as tracks, vertices, jets, EM clusters, muon candidates, shower max clusters, etc...
These objects - and the associated information - are then further scrutinized and eventually matched together to perform a full particle identification. The same strategy is pursued also at the trigger level (see sec. 3.5) where the reconstructed physics objects are employed in the online event filtering.
In this chapter details about the reconstruction algorithms relevant for this analysis will be provided. In particular we will address the photon identification and the jet reconstruction performed both at trigger level and at offline level.

4.2 Trigger Level Objects

We already mentioned in sec. 3.5 that a rudimentary object reconstruction is carried out during the online event filtering (trigger level). The achievable accuracy at this stage is limited by several factors as: (i) a partial account of the event information; (ii) a very short time constraint; (iii) the use of not-fully-calibrated quantities; (iv) the lack of vertex reconstruction, etc... Because of all these limitations only very loose requirements - whenever possible - are employed in the trigger object reconstruction in order to maintain an identification efficiency as high as possible.

4.3 Level 1 Objects

At L1 a generic electromagnetic object reconstruction is implemented to identify both electrons and photons. At this level the requirements are very loose and

just a tower [1] with an energy in the electromagnetic compartment greater than a certain threshold (we used 8 GeV) and with low activity in the hadronic one ($E_{HAD}/E_{CEM} \leqslant 0.125$) is enough to form a L1 EM cluster.
There is no jet identification at L1 but a scalar sum over all the towers ($\sum_{L1} E_T$) is performed at this stage. The very short time to carry out the calculation caused a high inaccuracy in the $\sum E_T$ evaluation (see Ch.6). Nevertheless it still gives an useful quantitative idea of the hadronic activity in the event.

4.4 Level 2 Objects

At this level more sophisticated reconstruction algorithms are implemented. We will focus only on two of them: (1) the EM cluster and (2) the calorimeter cluster ("*L2 jets*").

4.4.1 L2 EM Cluster

With this algorithm photon and electron candidates are selected. In the central calorimeter the energy deposited by these particles are generally confined in one or two towers. Starting from a "*seed tower*" (defined as a tower with $E^T_{CEM} > 8$ GeV) all the adjacent towers having at least 7.5 GeV of energy in the EM compartment ("*shoulder towers*") are grouped together to form a cluster and - if necessary - they are removed from the "seed tower" list. The process is then repeated for all the remaining "seed towers". For each cluster, besides the total energy, other two quantities are determined (fig. 4.1): (1) the isolation energy (`bestiso`), defined as the energy of the towers surrounding the EM cluster and (2) the isolation ratio (`bestrat`), defined as the isolation energy divided by the cluster energy. More details on the clusters algorithm can be found in reference [24].

4.4.2 L2 Calorimeter Cluster ("*L2 jets*")

A simple calorimeter cluster is implemented at L2 to help the jet identification. Each cluster starts from a "*seed tower*" with a threshold of 3 GeV (see fig 4.2). Then the cluster begins its expansion to the four orthogonal neighbors ("*shoulder towers*") of the seed tower if their energy are above the "shoulder" threshold of 0.125 GeV. After this first iteration, each new tower added become the starting point of a new "*shoulder towers*' expansion. The cluster will continue to expand until any new "*shoulder towers*" can be found.
By definition all the towers are unambiguously assigned to one cluster. In particular if a seed is found in a cluster initiated by another seed, it is removed from the seed list.
It should be noticed that in a crowded event - with a lot of towers above the shoulder threshold - the clusters may become very big. Moreover the seeds are

[1] a trigger tower (see sec. 3.5)

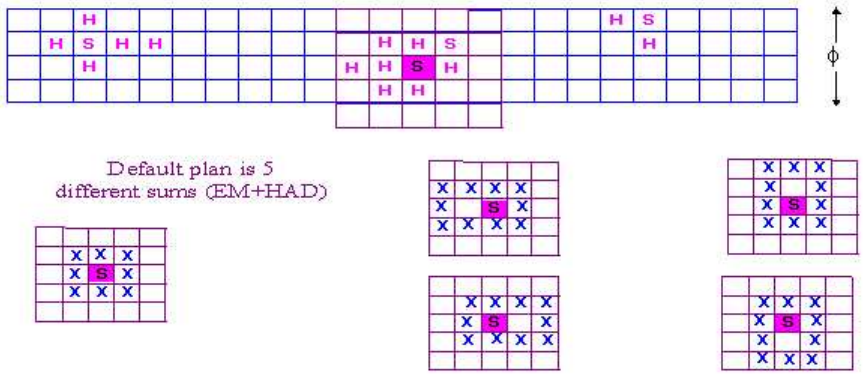


Figure 4.1: *The energy surrounding the EM cluster seed (S) is calculated in five different ways. The towers marked with a* X *are summed and five "surrounding energy" values are determined corresponding to the configuration depicted in the figure. The minimum among these five "surrounding energies" is defined as the isolation energy of the L2 EM cluster (*`bestiso`*). The configuration on the left is the only one used in Run I. In Run II the other four on the right were added to improve the performance of the algorithm.*

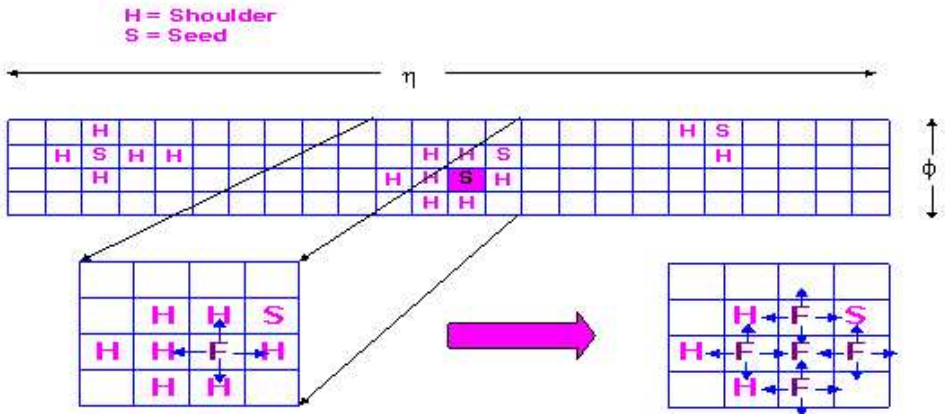


Figure 4.2: *The L2 calorimeter cluster algorithm starts with a trigger tower above threshold (*"seed tower"*). Then it begins to include (*F*) its orthogonal neighbors (*"shoulder towers"*). The process continues until no more* "shoulder towers" *can be found.*

not ordered in energy but the clustering starts from the seed with the lowest η, and within that η, from the lowest ϕ.
For each cluster found both the total electromagnetic and hadronic energy and

the seed towers η and ϕ coordinates are recorded along with the number of towers belong to the clusters.

4.5 Offline Objects

At the offline level the best and most accurate algorithms to reconstruct the physics objects are applied. The raw data collected by the detector are fully calibrated taking into account the slow response variation of the different subdetectors. At this stage the full vertex position and the beamline misplacement corrections are applied. Without any time constraint the reconstruction algorithms can be as much sophisticated as needed. In the following we outline the methods used to reconstruct the offline physics objects.

4.5.1 Vertex Reconstruction

The measurement of the vertex position along the z axis is essential to reconstruct correctly the kinematics of the event. Also, to properly take into account the extra energy coming from the secondary hard interactions, it is very important to measure the number of hard scatterings occurred in the same bunch crossing. The vertex algorithm uses the tracks reconstructed in the COT as "*seeds*". Tracks within 1 cm of a vertex seed are added together and the vertex location is calculated with an average of the track position weighted by their momenta. The vertex with the highest associated track P_T sum is defined as the primary vertex of the events.

4.5.2 CEM Clustering

The energy released by the electromagnetic particles (electrons and photons) in the calorimeter is usually contained in more than one tower. To identify these particles and their energy an EM calorimeter clustering is implemented. After a tower by tower calibration and a threshold cut of 100 MeV, all towers are sorted [2] in E_T. For the central calorimeter the clustering algorithm starts forming a list of towers with $E_T > 2$ GeV ("*seed towers*") [3]. Then for each "*seed tower*" a cluster is formed adding - if above threshold and if in the same tower - the two neighbor towers in η. In the end these added towers are removed from the list of the seed towers. It has to be noticed that no requirement to the hadronic over electromagnetic energy ratio is applied in this case.

4.5.3 Shower Max Clustering

When an electromagnetic shower reaches the central Shower Max Detector (see 3.3.1)several of its strips and wires generate a signal. To have a precise measurement of the EM shower position, a fixed window algorithm was implemented to

[2] The vertex is assumed at $z = 0$

[3] for the Plug calorimeter the strategy is similar but its geometry (see 3.3.2) requires a different definition of neighbor towers

group the Shower Max detector channels. The clustering strategy is identical for the strip and the wire chambers and it is very similar to that one implemented in Run I.
It proceeds as follow: (1) all the channels in the strip (wire) view are ordered in energy and a list of seeds is formed; (2) a cluster is made grouping eleven (seven) channels around the seed and the energy and the weighted position of the cluster is calculated; (3) the cluster profile is fitted with a *fixed* standard set of 10 GeV electron "standard" shower profiles measured during the 1985 test beam and a χ^2 is derived; (4) from the fit a new position is interactively set; (5) final corrections are applied to account for the small rotational and translation shift of the strips (wires) with respect to the standard CDF coordinate and for the differences between the particle energy and the test beam electron energy.
In Run II two different ways to form the seed list have been implemented. The first one (as in Run I) requires the seed channels to be above some threshold (150 MeV for the strips and 60 MeV for the wires). In this case at least two channels are required in the cluster to suppress noisy channel effects. The second possibility is a "*track-based*" cluster, where an extrapolated track position to the CES radius is used as a seed. In this case no further requirements are made on the strip (wire) energies. With these methods the EM object position can be measured with a precision of $\sim$1 mm.

4.5.4 Jet Clustering

The particles produced by the fragmentation of the parton generated in the hard $p\bar{p}$ interaction are distributed over a vast region around the parton original direction. The energy of the associated jet is spread across several calorimeter towers and a clustering algorithm is necessary to reconstruct the original jet energy and direction.
Various jet algorithms [25] have been developed in the past by many experiments and three of them are implemented and used to analyze the Run II CDF data [4] : a seeded cone-based algorithm (`JetClu`), a seedless cone-based algorithm (`MidPoint`) and a coneless-based algorithm (`kt algorithm`). In this analysis we employed the `JetClu` algorithm, a short description of which will be provided next.
The `JetClu` algorithm is an *iterative cone algorithm* which uses a cone with a fixed radius in η-ϕ space to define a jet. The clustering procedure consists of three steps: preclustering, clustering, and merging.

Preclustering

The clustering begins with creating a list of calorimeter towers with $E_T \geq E_T^{seed}$ which are used as seed towers for jets. The seed towers are stored in order of decreasing E_T [5]. In this phase the towers in the Plug calorimeter are grouped together to have the same segmentation as the central region. Preclusters are

[4] In Run I only the `JetClu` was implemented
[5] The transverse energy is calculated from the *reconstructed* z position of the primary vertex.

formed by combining adjacent seed towers within a cone of radius R in η-ϕ space. A seed tower is incorporated into a precluster if it is within the radius R of a seed tower with higher E_T which is already assigned to the precluster.

Clustering

Following the preclustering procedure, jet clustering is performed using the true tower segmentation. Jet clustering uses the E_T weighted centroid of a precluster $(\eta_{centroid}, \phi_{centroid})$,

$$\begin{aligned} \eta_{centroid} &= \frac{\sum_{i=1}^{n} E_T^i \eta^i}{\sum_{i=1}^{n} E_T^i}, \\ \phi_{centroid} &= \frac{\sum_{i=1}^{n} E_T^i \phi^i}{\sum_{i=1}^{n} E_T^i}, \end{aligned} \tag{4.1}$$

where the sums are carried out over all the seed towers in the precluster. The tower centroid (η^i, ϕ^i) is obtained by

$$\begin{aligned} \eta^i &= \frac{E_{T\,i}^{EM} \eta_i^{EM} + E_{T\,i}^{HA} \eta_i^{HA}}{E_T^i}, \\ \phi^i &= \frac{E_{T\,i}^{EM} \phi_i^{EM} + E_{T\,i}^{HA} \phi_i^{HA}}{E_T^i}, \end{aligned} \tag{4.2}$$

where $E_{T\,i}^{EM}$ and $E_{T\,i}^{HA}$ are transverse energies deposited in the electromagnetic (EM) and hadronic (HA) parts of a calorimeter tower with index i. $(\eta_i^{EM}, \phi_i^{EM})$ and $(\eta_i^{HA}, \phi_i^{HA})$ are the centroids of the electromagnetic and hadronic components of tower i, defined by a vector pointing from the event vertex to the center of the calorimeter tower (calculated at the depth that corresponds to shower maximum). A cone of radius R in η-ϕ space is created around the centroid of a cluster. Then, all the towers with $E_T \geq 100$ MeV are incorporated into the cluster if the towers are within the cone. A new cluster centroid is determined from the towers within the cone using an E_T weighted centroid, and a new cone is created using the new cluster centroid. This process continues until the tower list remains unchanged.

Merging

At the stage of clustering, some towers may be shared by more than one cluster. If towers of one cluster are completely contained within another cluster, the smaller cluster is dropped. If two clusters partially overlap, an overlap fraction is computed by summing the E_T of the shared towers and dividing it by the E_T of the smaller cluster. If the fraction is above a cutoff value of 0.75, then the two clusters are merged. If the fraction is less than the cut threshold, the clusters are kept unchanged and the shared towers are assigned to the nearest cluster in η-ϕ space. After the towers are assigned uniquely to

clusters, the centroid computation and tower shuffling are repeated until the tower lists stay unchanged. Finally a cluster four-vector is defined by summing over the four-vectors [6] of the towers belonging to the cluster.

4.6 Jet Energy Correction

The momentum four-vector of reconstructed jets generally differs from the energy of the initial parton from which it has been generated. This results both from the jet particle finite resolution response of the detector and from the intrinsic limitation of the jet algorithms to identify all (and only) the particle from the parton fragmentation. Therefore to retrieve the four-momentum of the initial parton, proper corrections, accounting both for detector effects and physics effects - need to be applied. The jet corrections [7] are composed of five different stages (or "*level*") applied in sequence, with the output of one level being the input of the following and each one accommodating different effects. The jet corrections depend upon the clustering cone radius R, the momentum measured in the cone P_T, the detector pseudorapidity of the cluster centroid η, the number of primary vertexes N_{vtx} and the Run number of the event `RunNumber`. The path to determine the final corrected P_T of jet (P_T^{corr}) is the following (level-7 means $full$ corrections) [8]:

$$\begin{aligned}
\text{Level 1:} \quad & P_T^{rel} = f_\eta(R, P_T, \eta, \texttt{RunNumber}) \times P_T^{raw} \\
\text{Level 4:} \quad & P_T^{MI} = P_T^{rel} - MI(R) \cdot (N_{vtx} - 1) \\
\text{Level 5:} \quad & P_T^{abs} = f_{jes}(R, P_T) \times P_T^{MI} \\
\text{Level 6:} \quad & P_T^{UE} = P_T^{abs} - UE(R) \\
\text{Level 7:} \quad & P_T^{corr} = P_T^{abs} \times OOC(R, P_T)
\end{aligned}$$

A short description of each correction level and of the method used to determine them will be outlined in the following paragraphs [26].

4.6.1 Relative Correction: $f_\eta(R, P_T, \eta, \texttt{RunNumber})$

To account for the non-uniform response of different calorimeters and for the undetected energy falling in uninstrumented regions of the detector ("*cracks*"), a correction is applied to make the calorimeter response to jet energies uniform in η. This process results in correcting jets back to an equivalent jet in the eta range $0.2 \leqslant |\eta| \leqslant 0.6$ where the calorimeter response is flat and non-linearities

[6] The tower four-vector is built massless but the cluster four-vector is in general not massless.

[7] Up to now only the corrections for jets made with the `JetClu` algorithm have been derived

[8] Level-2 and level-3 were used in a early version of the jet corrections. Once those levels have been removed, for code back-compatibility the levels have not been renamed.

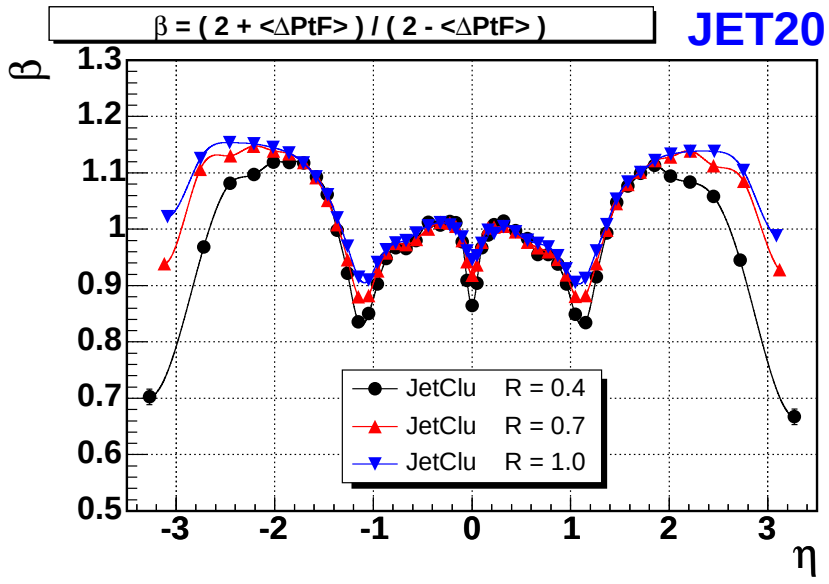


Figure 4.3: *Relative correction factor f_η as a function of η of the probe jet for different cone size. Jets from dijet events with any* `RunNumber` *and with* $25 \leqslant P_T^{raw} \leqslant 55$ *are used in this plot.*

are well understood from test-beam measurements.
To determine the relative corrections, the p_T balancing of the first two leading jets (Δp_T) in dijet events [9] is used. One of the two jets - *trigger jet* - is required to be in the central region ($0.2 \leqslant |\eta^{trigger}| \leqslant 0.6$) whereas no restriction is applied on the position of the other jet - *probe jet*.

Since for an ideally uniform detector we have (on average) $p_T^{probe} \cong p_T^{trigger}$, a relative correction factor f_η is extracted from the Δp_T observed in the data (see fig. 4.3).

The f_η factor depends on the jet η, the jet P_T and the jet cone radius R, along with a small time dependence (`RunNumber`). The uncertainties on f_η come from the accuracy of the procedure, the event selection, the extrapolation of the correction in region not covered by measurement (low P_T^{jet}), the inaccuracy of the MC modeling of dijet events and from the process-dependence relative response of the jet energy.

[9] In Run I the missing E_T projection fraction was used to calculate the relative correction. A comparison of the two methods found the balancing method less sensitive to the difference between the central and the plug calorimeter to the fluctuation of the energy outside the clustering cone.

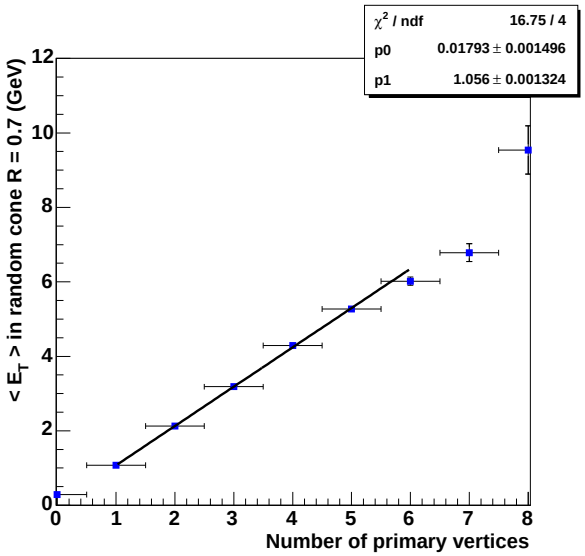


Figure 4.4: *Average energy in a random cone with R=0.7 in minimum bias data as a function of the number of vertices. The slope of the fit gives the multiple interaction corrections $MI(R)$. The value at $N_{vtx} = 1$ is used to define the underlying event correction UE(R).*

4.6.2 Multiple $p\bar{p}$ interactions: $MI(R)$

When multiple $p\bar{p}$ interactions occur in the same bunch crossing, some particles coming from the extra min-bias events can fall into the jet clustering cone. The energy released from these particles - not coming from the parton fragmentation process - have to be subtracted from the jet energy. This correction has been derived measuring, as a function of N_{vtx}, the transverse energy in a random cone in min-bias data. The additional energy per vertex (slope in fig. 4.4) is found to be 0.356 (1.056, 2.153) for a cone size of 0.4 (0.7, 1.0). The systematic uncertainties has been estimated using different topology samples and evaluated at $\sim 15\%$ level for all cone sizes.

4.6.3 Absolute Jet Energy Scale: $f_{jes}(R, P_T)$

After the relative corrections all the jets have the same scale as the jets in the central region. These jets have to be corrected for the non-linear response of the central calorimeter and for the energy loss in the cracks or by invisible particles (low p_t particles that never reach the calorimeter).
Montecarlo dijet events are used to determine the absolute jet energy scale. Both the standard calorimeter clustering algorithm and a hadron-level jet clustering are applied. The latter uses the same algorithm of the former but applied

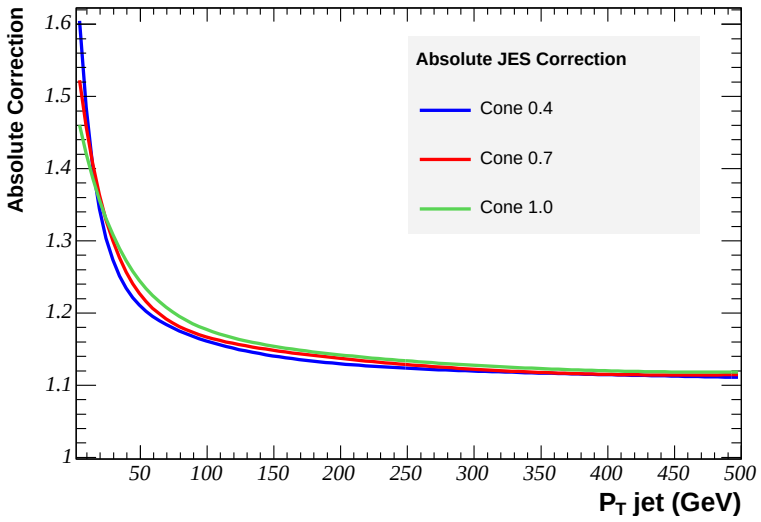


Figure 4.5: *Absolute jet energy scale correction f_{jes} as a function of P_T and for different cone radius R.*

on the MC list of the stable particles after the hadronization (HEPG particles) instead than on calorimeter towers. It has to be noticed that the hadron-level jets include particles coming from the parton hadronization as well as particles from spectator interactions ("underlying event") and from initial state radiation (ISR) which may fall within the jet cone. After the matching between the calorimeter-level jets and the hadron-level jets ($\Delta R(j^{cal}, j^{had})$ <0.1), a correction factor $f_{jes}(R, P_T) \equiv P_T^{had}/P_T^{cal}$ is derived as a function of P_T^{cal} and cone radius R (fig. 4.5).
The uncertainties on absolute corrections arise mainly from the inaccuracy of the detector simulation and the MC fragmentation model.

4.6.4 Underlying Events: $UE(R)$

The energy inside the clustering cone produced by remnants of the $p\bar{p}$ system not involved in the hard scattering ("*spectator interactions*") has to be subtracted from the jet energy. This correction may depend upon the process under study. For simplicity the spectator interactions are assumed well modeled by minimum bias events and this correction has been derived using the method described in sec. 4.6.2 for the case of only one vertex. The underlying event energy is found [10] to be 0.6 (1.6, 3.2) GeV for a 0.4 (0.7, 1.0) cone radius.

[10] An absolute correction factor of 1.6 has to be applied to the unclustered energy since this correction come after the absolute correction.

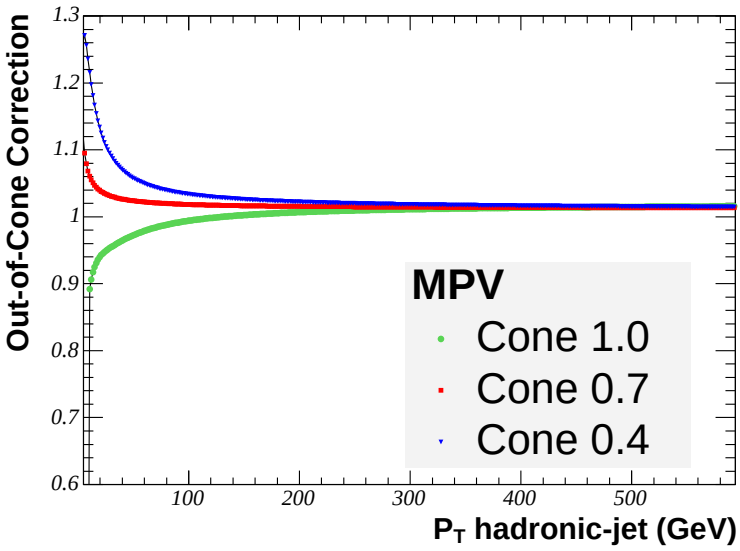


Figure 4.6: *Out of Cone correction $OOC(R, P_T)$ as a function of P_T for different cone size. By definition this correction includes the underling event subtraction. For jets with cone radius of 1.0 the extra energy from underlying events outcomes the energy lost outside the cone requiring an overall subtraction of energy ($OOC < 1$).*

4.6.5 Out of Cone: $OOC(R, P_T)$

The jet clustering may not include all the energy coming from the initiating parton. Indeed, some particles produced during the fragmentation process can fall outside the clustering cone. This energy have to be estimated and added to the energy of the jet.
This correction is determined using dijet MC sample. Hadron-level jets in the central region are matched to partons and the correction factor $OOC(R, P_T) \equiv P_T^{part}/P_T^{had}$ is obtained with the same parametrization strategy used for the absolute corrections (fig. 4.6). As mentioned in sec. 4.6.3 the hadron-level jets include the underlying events. Therefore the OOC correction contains implicitly the underlying event subtraction contribution and therefore it has to be applied after the absolute corrections. This is the reason why for large cone we can have $P_T^{had} > P_T^{part}$ making the OOC a subtracting correction overall. The systematic uncertainties are evaluated looking at the calorimeter energy deposited inside an annulus from the cone boundary to R=1.3 on data and MC γ+jet events. Their difference is taken as a systematics uncertainty.
The total jet energy correction systematic uncertainty along with the amount of each different contribution is reported in fig. 4.7 as a function of P_T^{corr} for central jets.

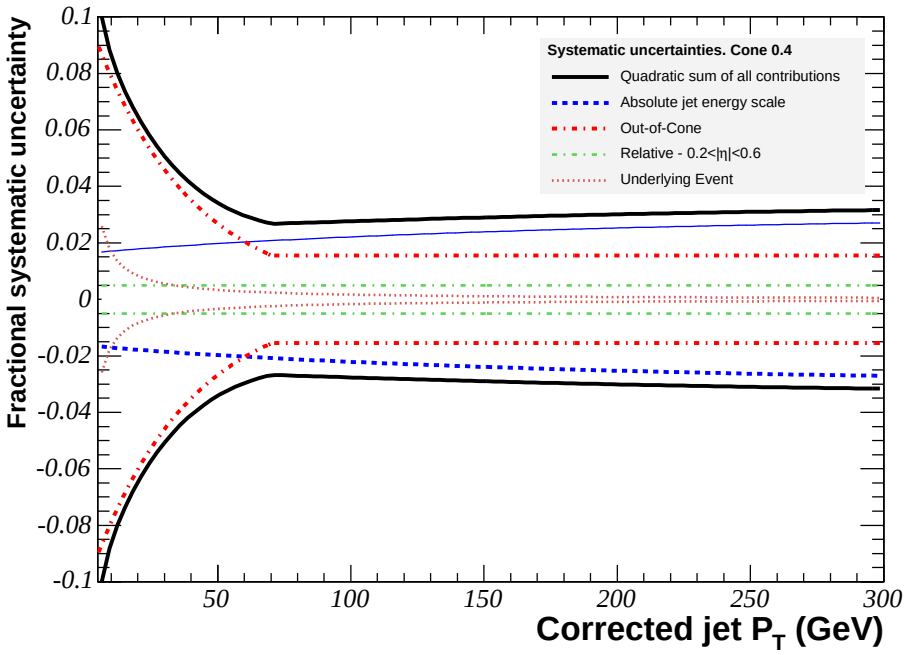


Figure 4.7: *Different contributions and total jet correction systematic uncertainty as a function of the corrected jet P_T in the region $0.2 \leqslant |\eta^{jet}| \leqslant 0.6$ for jet with cone radius 0.4*

4.7 Photon Identification

The identification of isolated low p_t photon is one of the key ingredient for this analysis. Indeed our search for the W and Z hadron decay strategy is to select events with a photon first and then look for presence of a heavy vector boson on these events. The photon event selection is based on a sequence of quality cuts over quantities derived from the object reconstruction process described earlier in the chapter. First of all the presence of an EM cluster object is required for a photon identification. Further restrictions are then applied to reduce the background coming from jets, meson decays and cosmic ray muons. This is done in two stages. On the first step some straight cuts based on several different photon ID variables suitable to discriminate photons from background are applied and a subsample of *photon candidates*" is selected. Then a statistical subtraction of the neutral meson multiple photon decay contamination is performed. In this second step no cuts are set, but the fraction of prompt photons is evaluated on a statistical basis.

4.7.1 Photon Candidates Selection

Photon candidates are selected with the following cuts:

- **Fiducial:** This basic cut requires a photon to be located in the region of the central calorimeter covered by the CES and CPR detectors. Therefore the photon position has to be within 17.5 cm of the CES chamber center in the azimuthal direction (perpendicular to the wires) and within the active region of the strip chamber (14< $|z|$ <217 cm).

- **Isolation:** Very high collimated jets with a large electromagnetic component can be misidentified as prompt photons. Requiring a very low activity (E_T < 1.0 GeV) in a cone of radius 0.4 around the photon suppresses this background.

- **Tracks:** Electrons and other charged particles can be eliminated requiring a veto on EM clusters associated with tracks. Thus, no reconstructed COT tracks pointing at the CPR chamber (or at the tower the photon candidates is located) area allowed.

- **Extra-CES clusters:** The single and multiple meson background is reduced by requiring no other extra CES clusters - with an energy above 1 GeV - to be associated to the EM object.

- **CES χ^2:** The shower shape measured by the CES is one of the most important variable to select prompt photon. At this point only a cut $\chi^2 < 20$ is applied for photon candidates but the profile will play an important role on the second step of the prompt photon selection.

- **Hadronic Energy:** To suppress hadronic background a cut on the ratio of the electromagnetic and the hadronic energy of the cluster is set. It is required that $E_{HAD}/E_{CEM} < 0.055 + 0.00045 \cdot E^{\gamma}(GeV)$. The scaling with the energy of the photon allows some punch-trough of more energetic photons.

- **Missing E_T:** Cosmic ray muons can radiate a photon in the CEM and fake a photon signal. The characteristic signature for this process is that there is no jet on the other side of the event. The cosmic ray background is eliminated by cutting on the amount of missing energy divided by the cluster E_T (or photon E_T) [27]

All the cuts described above are summarized in Table. 4.1.

4.8 Photon Background Subtraction

A major background left after the selection cuts presented in the previous paragraph is the multiple photon decay of neutral mesons such as π^0, η. To evaluate the fraction of prompt photons in the selected photon candidate sample, two

Variable	Cuts
Detector region	Central Only
Fiducial	$\lvert X_{CES}\rvert$ <17.5 cm 14< $\lvert Z_{CES}\rvert$ <217 cm
Isolation	E_T(Cone 0.4) < 1 GeV
Extra-CES Clusters	$E_T < 1$ GeV
Matched COT Tracks	None Allowed
CES $((\chi_{strip} + \chi_{wire})/2$	< 20
Hadronic Energy	$E_{HAD}/E_{CEM} < 0.055 + 0.00045 \cdot E^{\gamma}(GeV)$
Missing E_T	$\sum \not\!\!E_T/E_T^{\gamma} < 0.8$

Table 4.1: *Photon candidates selection cuts*

techniques have been developed, similar in strategy but using different event information: the profile (CES) method and the conversion (CPR) method.
The profile method exploits the differences between the CES shower shape generated by a single photon and that one produced by multiple photons. The shower formed by two nearby photons from meson decays is more likely to be wider (with larger χ^2) than those formed by single photon. The CES shape of test beam electrons is used as single photon shape reference.
The CES method is very effective at low P_T as multiple photon shower is typically widely separated. At high P_T instead, since the multiple photons from meson decays are very collimated, the distinction between them and the prompt photons is impossible.
The CPR method looks at the fraction of photons that convert in the material located in front of the CPR detector. The conversion probability [11] of multiple photons is greater than 80% while for a single photon is about 60%. Furthermore this method can be applied in the all P_T range although the CES method is a little more accurate (when it can be used). A description of the general strategy and details about the two methods follow.

4.8.1 Statistic Background Subtraction

Suppose to have a data sample made of a mixture of signal and background events, an observable O and a cut-off - or threshold - for this observable, Λ. The total number of events are N_{tot} with N_p above Λ . The total number of signal and background events are N_s and N_b respectively ($N_{tot} = N_s + N_b$). The fraction of signal, background and data events above Λ are ϵ_s, ϵ_b and ϵ. We have thus:

$$\begin{cases} \epsilon_s \cdot N_s + \epsilon_b \cdot N_b = N_p \\ (1-\epsilon_s) \cdot N_s + (1-\epsilon_b) \cdot N_b = N_{tot} - N_p \end{cases}$$

[11] of at least one conversion

After some simplification we get

$$N_s = N_{tot} \cdot \frac{\epsilon - \epsilon_b}{\epsilon_s - \epsilon_b}$$

This formula allows to calculate the *total* number of signal in the data if the ϵ_s and ϵ_b are known (for the data we have $\epsilon = N_p/N_{tot}$). It should be noticed that although in principle Λ is arbitrary, it is usually determined in order to minimize the systematic uncertainties (both ϵ_s and ϵ_b come with errors). As a rule of thumb it should give a good separation between the signal and the background in order to maximize $\epsilon_s - \epsilon_b$ [12].
For the conversion method the observable is the charge observed in the CPR with a $\Lambda_{CPR} = 500\ fC$ while for the profile method the observable is the CES χ^2 with $\Lambda_{CES} = 4$.

4.8.2 The Conversion (CPR) Method

The conversion method uses the fact that one of the two photons from a neutral meson decay will convert in the coil material in front of the calorimeter with higher probability than a single photon. In this case ϵ is the fraction of photon candidates producing a pulse height in the CPR - within a 66 mrad "*window*" (5 CPR channels) around the photon direction [13] - greater than the one produced by a minimum ionizing particle (500 fC).
Generally, the probability for a single photon to pass without conversion through X_0 radiation length of material is $e^{-\frac{7}{9}X_0}$. Thus the conversion probability for direct and secondary photons can be obtained from [29]:

$$\begin{cases} P_\gamma = 1 - e^{-\frac{7}{9}X_0} \\ P_{\pi^0} = 1 - (1 - P_\gamma)^2 \end{cases}$$

However there a number of corrections which need to be made as discussed below. The first thing to correct for is the effective number of photons detected within the CPR *"window"*, N_γ [28]. For low energy π^0 's and η's the separation between the two photons is large enough that only one photon is in the *"window"*. For other multiple photon decays of the η and K_s this is not longer true. N_γ is also P_T dependent because the photons from meson decays are more collimated at high P_T. Beyond some energy virtually all the photons are in the "*window*" all the time. The conversion probability is therefore generalized as

$$P = 1 - (e^{-\frac{7}{9}X_0})^{N_\gamma(P_T)} = 1 - e^{-\frac{7}{9}X_0 N_\gamma(P_T)}$$

where for a single direct photon we have $N_\gamma(P_T) = 1$, for all P_T. Other corrections come from underlying events and from dead CPR channels. The probability to *not* have a conversion $(1 - P)$ has to be multiplied by the probability that the underlying event does not produce a CPR hit, f_{UE}. Then the

[12] But this is not the only criteria. For a more complete discussion of the problem see [28]
[13] For reference the minimum separation of the photon from a 25 GeV π^0 is 11 mrad

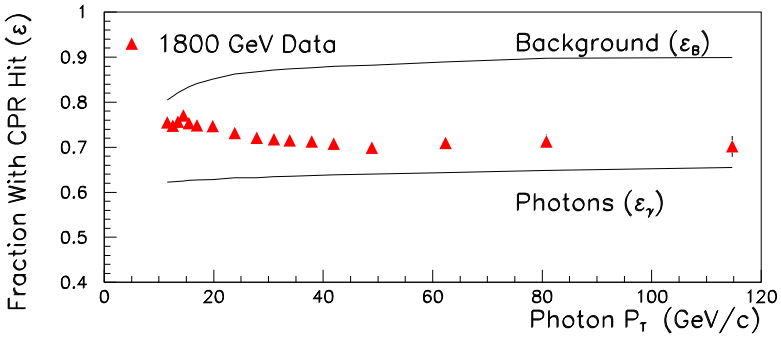


Figure 4.8: *CPR efficiency for the signal (ϵ_s), the background (ϵ_b) and the data (ϵ) as measured with Run I data. The P_T dependence of ϵ_b arise from the correction to the number of photons falling in the CPR "window" region. The same CPR efficiencies are used in Run II data.*

corrected conversion probability has to be multiplied by the fraction of working channels $(1 - f_{dead})$. Thus

$$P = [1 - (1 - P) \cdot f_{UE}] \cdot (1 - f_{dead})$$

where f_{dead} is the fraction of CPR dead channels (3.5%). In Run I selecting a random cone of radius 0.4 in minimum bias data, the chance for a hit in a five channel window was found to be 4.9%, with a dependence on the number of vertexes. In Run II isolated muon candidates from a $W \to \mu\nu$ sample have been used to study the CPR response to single minimum-ionizing particles [30]. The underlying event activity of the W and photon sample is assumed to be very similar. This study results in a slightly different f_{UE} with respect to Run I. We decided to adopt the Run I value for the analysis reported in this dissertation and include the Run I/Run II difference in the systematics.
Finally, back-splash photons from the electromagnetic shower in the calorimeter can convert or Compton scatter and give a hit in the CPR. This effect was studied using the GEANT simulation and an additional correction was derived [28]. It has the form

$$P_{BS} = (1 - P_{UE}) \cdot f_{bs}$$

with a back-splash factor $f_{bs} = 0.00074 \cdot P_T / \sin\theta$.
The accounting of the material was completely redone with Run II data [31] since the detector significantly changed. The total expected thickness of material at normal (90 degree) incidence was estimated at $(1.072 \pm 0.018)\ X_0$.
The CPR method efficiencies for the data, signal and background are shown in fig. 4.8. The photon curve does not have a P_T dependence whereas for the

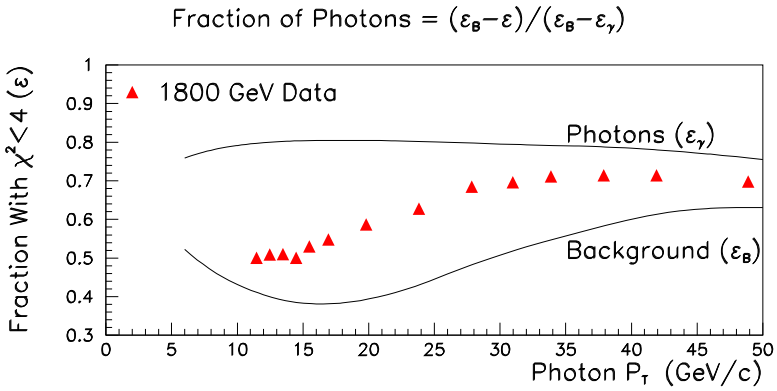
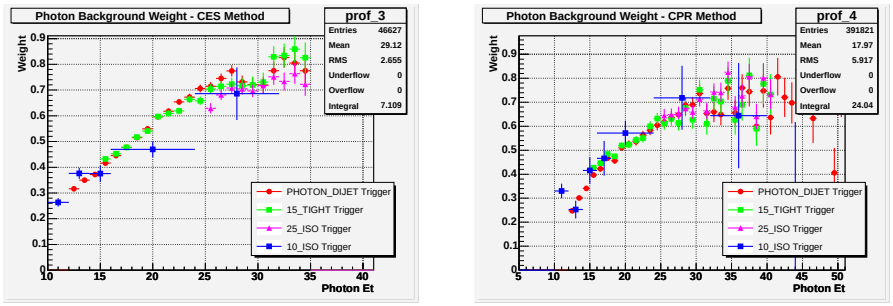


Figure 4.9: *CES efficiency for the signal (ϵ_s), the background (ϵ_b) and the data (ϵ) as measured with Run I data. The same CES efficiencies scaled by a 5% factor are used in Run II data.*

background the P_T dependence is very small only at high P_T when all secondary photons are within the CPR.



(a) CES Method (b) CPR Method

Figure 4.10: *Prompt photon fraction as measured in Run II with the CES method (left) and CPR method (right) photon background subtraction as a function of of the P_T^γ. Inclusive photon samples triggered at different threshold are used to cover the all photon P_T spectrum (A photon trigger description is provided in sec. 6.1). The errors are statistical only.*

4.8.3 The Profile (CES) Method

For this method the efficiency ϵ is defined as the fraction of photon candidates with $\chi^2 < 4$ (the description of the CES clustering and χ^2 definition was provided in sec. 4.5.3). In Run I ϵ_s and ϵ_b have been measured from data and from simulation. The fig. 4.9 presents the CES efficiency as a function of the photon P_T. The data (points) lie between the two curves of the expected signal and background efficiency. In Run II the CES $\chi^2 < 4$ efficiency was measured reconstructing the $\eta \to \gamma\gamma$ peak - for low E_T photons - and using electrons from W and Z - to cover the high E_T region. Both methods measured an efficiency 5% lower than in Run I. This degradation was confirmed by comparing the CPR method and the CES method tuned with the Run I efficiency. We decided to adopt the lower CES Run II efficiency for the analysis reported in this dissertation and - as for the CPR method - include the Run I/Run II difference in the systematics.
The fraction of prompt photons measured in with the Run II data and determined with the CES and CPR methods are reported in fig. 4.10 as a function of the P_T^γ.

Chapter 5

Signal Expectation

5.1 Introduction

The purpose of the analysis reported in this dissertation is the search of the hadronic decays of the W and Z bosons produced in association with a photon. Theoretical calculations of $W\gamma$ and $Z\gamma$ production in $p\bar{p}$ collisions appear copiously in literature (see for example [32] and reference within). Such interest arises from the sensitivity of these processes to the three-linear gauge boson couplings whose measurement provides a crucial test of the non-Abelian nature of the Standard Model (SM). Indeed two of such three boson couplings - $ZZ\gamma$ and the $Z\gamma\gamma$ - are predicted to be zero whereas the $WW\gamma$ coupling is completely fixed by the gauge theory structure of the Standard Model (at the tree level). Thus, any deviation from the SM prediction might be interpreted as a contribution of a beyond SM non-gauge theory.

At the Born level (LO) the reaction $p\bar{p} \to W(Z)\gamma \to q\bar{q}\gamma$ proceeds via the Feynman diagrams shown in fig. 5.1(a). Technically, if a finite-width $W(Z)$ propagator is taken into account, in order to preserve electromagnetic gauge invariance the reaction $p\bar{p} \to W(Z) \to q\bar{q}\gamma$ (whose diagrams are shown fig. 5.1(b)) have to be included in the calculation (these radiative $W(Z)$ decays are also known as "*inner bremsstrahlung*"). In principle, because of the interference terms the two contributions can no longer be distinguished. However - as it will be shown later - an excellent separation between the two contributions can be achieved imposing some suitable cuts on the final state observables. This possibility is particularly important for this analysis because the mass reconstruction of the $q\bar{q}$ system gives the right $W(Z)$ boson mass value only in the direct $W(Z)\gamma$ production case. In the inner bremsstrahlung mechanism instead, part of the $W(Z)$ energy is carried away by the radiated photon and the $W(Z)$ mass cannot be reconstructed with only the hadronic final state informations. For this reason in this analysis we define as signal *only* the process $p\bar{p} \to W(Z)\gamma \to q\bar{q}\gamma$ whereas the inner bremsstrahlung is considered as part of the background.
The LO theoretical calculations of the $W(Z)\gamma$ process are encoded in several

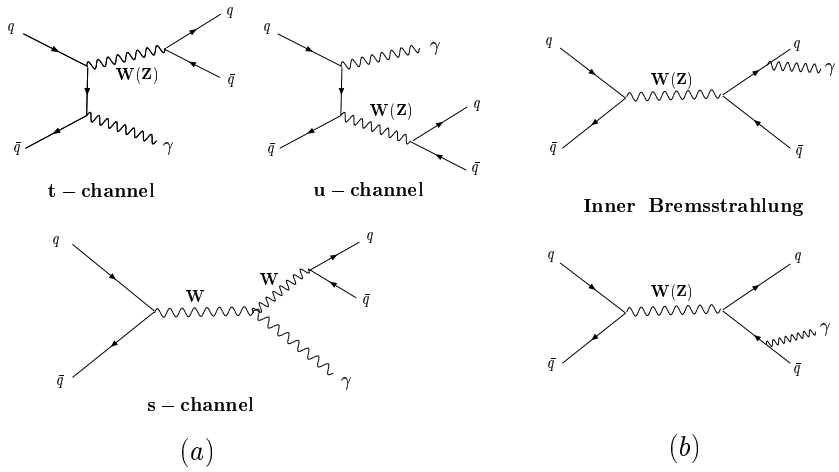


Figure 5.1: *(a) Feynman diagrams at the tree level for the process* $q\bar{q} \to W(Z)\gamma \to q\bar{q}\gamma$. *The s-channel for the* $Z\gamma$ *production is forbidden in the Standard Model. (b) Radiative W(Z) decays diagrams. A bremsstrahlung photon is emitted by one of the two quarks from the W(Z) decay. These two diagrams are part of the* background *for this analysis.*

matrix element (ME) Monte Carlo generator programs. Generally they deal only with the parton level calculation of the hard interaction providing the process cross section - after a phase space integration - along with the final state parton four-vectors. The complete event generation - including the parton shower, the hadronization and the modeling of the initial state gluon radiation - is carried out by other general-purpose Monte Carlo like PYTHIA [33] or HERWIG [34]. Of course PYTHIA (or HERWIG) itself can be used as a ME Monte Carlo generator.

5.2 Monte Carlo Signal Generation

To generate our $p\bar{p} \to W(Z)\gamma \to q\bar{q}\gamma$ signal sample we decided to use PYTHIA (version 6.203). It is important to stress that PYTHIA *does not* properly include the inner Bremsstrahlung diagrams (it uses the narrow $W(Z)$-width approximation). As far as the signal generation is concerned, this is not a problem because - as we mentioned before - the Inner Bremsstrahlung is not part of this analysis signal. However, the lack of the interference terms in the calculation may in principle leads to some inaccuracy in the direct $W(Z)\gamma$ Monte Carlo prediction. To address this issue we performed a parton level comparison (sec. 5.2.2) between PYTHIA and MadGraph, a ME Monte Carlo generator

that employs all the $W(Z)\gamma$ diagrams in its calculation.

5.2.1 PYTHIA Prediction

The PYTHIA Monte Carlo generator is a "general-purpose" Monte Carlo that can simulate a wide range of reactions occurring in a $p\bar{p}$ collision (and other kind of collisions). Besides the perturbative calculation of the hard interaction, PYTHIA is able to model the parton shower, the initial gluon radiation and the hadronization process along with a tunable underlying event contribution, providing a complete description of the event at the observable level.
For the $W\gamma$ and $Z\gamma$ production PYTHIA includes with the tree level u,t and s-channel diagrams and their interferences but not the inner bremsstrahlung diagrams [1]. The main advantage of using PYTHIA is the possibility to generate *all* the signal subprocesses simultaneously avoiding the annoyance of weighting the events from every subprocess in accordance to their relative cross sections. Considering that there are 5 possible decay channels for the W and 6 for the Z, the signal generation with PYTHIA is highly simplified.
For the generation of our signal sample we set the factorization scale, q, equal to the incoming quark momentum transfer, $\sqrt{\hat{s}}$. The initial parton momentum distributions are determined using the CTEQ5L [36] LO parton distribution function (PDF). Because of the infrared divergences when the photon energy approaches zero, we set a minimum photon p_t cut of 6 GeV. No further cuts on the final partons were set at the generation level.
Finally the generated events were filtered using loose kinematic and geometrical limits close to the analysis selection criteria. All the events without a photon with a $P_T > 10$ GeV and $|\eta| < 1.2$ were rejected. In this way we saved a huge amount of CPU time avoiding the full detector simulation of events we would have discarded during the analysis selection. After the post-generation filter the remaining signal events were processed by the CDF Detector Simulation program. It is a GEANT based Monte Carlo that simulates the interaction between the produced particles and the detector material. All the details of each subdetector components are included in the program (geometry, active and passive material, etc...) and effects induced by multiple scattering, bremsstrahlung, secondary particle generation are properly modeled. Eventually, the output of the simulation is handled by the DAQ in exactly the same way of the real collision data, using the same offline reconstruction code and storing the output in the same format.
We generated 1.5M events and, after the post-generation filter, we ended up with 485.155 events. The corresponding total LO PYTHIA cross section is reported in Table. 5.1 along with the individual channel cross sections.

[1] It has to be noticed that the soft photon emission is part of the parton shower process and therefore - in principle - it can be introduced in PYTHIA as well. However, even under these circumstances, the correct interference terms would not be present.

PYTHIA Cross Section Prediction		
	Process	*Cross Section* (pb)
Generation cuts: $P_T^\gamma > 6$ GeV	$W^+ \to u\bar{d}$	1.59
	$W^- \to d\bar{u}$	1.59
	$W^+ \to c\bar{s}$	1.59
	$W^- \to s\bar{c}$	1.59
	$W \to q\bar{q}$	6.36
Post-Generation Filter: $P_T^\gamma > 10$ GeV $\lvert\eta^\gamma\rvert < 1.2$	$Z^0 \to u\bar{u}$	1.27
	$Z^0 \to c\bar{c}$	1.25
	$Z^0 \to d\bar{d}$	1.62
	$Z^0 \to s\bar{s}$	1.62
	$Z^0 \to b\bar{b}$	1.62
	$Z^0 \to q\bar{q}$	7.4
Total Cross Section: 13.76 pb		

Table 5.1: *PYTHIA cross section prediction for the processes $p\bar{p} \to W\gamma$ and $p\bar{p} \to Z\gamma$. The cross section for each boson channel decay is reported.*

5.2.2 MadGraph Prediction

To evaluate the accuracy of the PYTHIA cross section predictions, we generated another sample of $W\gamma$ and $Z\gamma$ events with a different Monte Carlo. This time we used the MadGraph Monte Carlo generator [35], a tree level matrix element calculator that properly *includes* the radiative W(Z) boson decay in the $W(Z)\gamma$ process generation. After the integration over the phase space MadGraph provides the cross section prediction along with all the relevant information (momenta, spin, particle id's,...) of the final partons which can eventually be passed to a shower Monte Carlo program.
As in PYTHIA, at the beginning we applied very loose generation cuts - $P_T^\gamma > 6$ GeV and $|\eta^\gamma| < 3$ - and then we filtered the final sample requiring the presence of a photon with $P_T > 10$ GeV and $|\eta|$ <1.2. However, the presence of photon radiative W(Z) decays causes divergences in the cross section calculation when the photon is collinear to the emitting quark. To remove these collinear divergences we required (at generation level) the distance between the photon and the quarks in the η-ϕ space to be greater than 0.4 ($\Delta R(\gamma, q) > 0.4$).
Since in MadGraph the user has to specify the flavor of the final state particles, we had to generate separately all the possible decay modes of the $p\bar{p} \to W(Z)\gamma$ process [2]. The parton distribution functions and the factorization scale were set, as in PYTHIA, to CTEQ5L and $\sqrt{\hat{s}}$ respectively.

[2] Actually MadGraph can accept some set of particles as a input. In particular a *proton* can be provided as initial particle, in which case MadGraph will sum over the quarks and gluons initial state. Similarly a generic *jet* can be provided as final particle but the sum in this case doesn't include heavy quarks (b and t).

The first issue we addressed with the new sample was the separability of the direct production from the inner bremsstrahlung. The di-quark invariant mass distribution (for the $Z\gamma$ sample) along with a di-quark invariant mass versus the three-body invariant mass scatter plot are shown in fig. 5.2. It is clear for these plots that the two contributions can practically be separated with very high accuracy applying a cut on the di-quark invariant mass. It has to be noticed that this good separability is a consequence of the relatively high threshold for the photon momentum ($P_T^{\gamma}>$ 10 GeV). For softer photons the two contributions cannot be longer effectively separated.
The MadGraph cross section predictions for the *direct only* subprocesses are summarized in Table. 5.2. For completeness we reported as well the fraction of inner bremsstrahlung events for each channel.

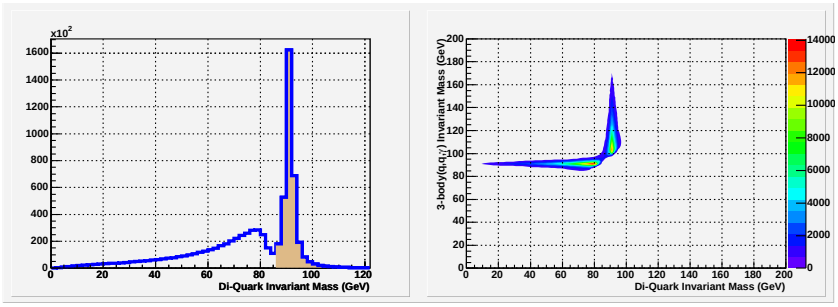


Figure 5.2: *Left: the $q\bar{q}$ reconstructed invariant mass distribution of $p\bar{p} \to Z\gamma \to q\bar{q}\gamma$ (yellow shadow) and $p\bar{p} \to Z \to q\bar{q}\gamma$ events generated with* MadGraph *is shown. Right: the same quantity is plotted vs. the 3-body (q,q,γ) invariant mass. A clear separation between the direct $Z\gamma$ production and the inner bremsstrahlung contribution is noticeable. The $W\gamma$ events present the same features.*

Since we used the same selection cuts, the PYTHIA and MadGraph cross section predictions can be directly compared. The total cross sections agree within 11%, with a better agreement for the $\sigma_{Z\gamma}$ (6%) than for the $\sigma_{W\gamma}$ (25%).

To further check the consistency between PYTHIA and MadGraph signal predictions we made a comparison of some final state observable distributions. In particular we compared the P_T distribution of the three final partons, their η and the distance between each other, ΔR. The comparison was made both for $Z\gamma$ events (fig. 5.3) and for $W\gamma$ events (fig. 5.4). A very good agreement between all the normalized distributions is observed. It is clear that because the simulation processes after the parton level generation are exactly the same for PYTHIA and MadGraph (parton shower, hadronization, etc...), the consistency of the final parton kinematics distributions implies a full agreement between the two Monte Carlo as far as the signal event simulation is concerned.

MadGraph **Cross Section Prediction**			
	Process	*Cross Section* (pb)	*Bremss. Fraction*
Generation cuts:	$W^+ \to u\bar{d}$	2.01	63%
$P_T^\gamma > 6$ GeV	$W^- \to d\bar{u}$	1.96	62%
$\lvert\eta^\gamma\rvert < 3$	$W^+ \to c\bar{s}$	2.24	57%
$\Delta R(\gamma, q) > 0.4$	$W^- \to s\bar{c}$	2.25	57%
	$W \to q\bar{q}$	8.46	60%
Post-Generation Filter:	$Z^0 \to u\bar{u}$	1.16	48%
$P_T^\gamma > 10$ GeV	$Z^0 \to c\bar{c}$	1.14	48%
$\lvert\eta^\gamma\rvert < 1.2$	$Z^0 \to d\bar{d}$	1.53	57%
$W\gamma$: $M_{q\bar{q}} > 74$ GeV	$Z^0 \to s\bar{s}$	1.57	56%
$Z\gamma$: $M_{q\bar{q}} > 85$ GeV	$Z^0 \to b\bar{b}$	1.56	54%
	$Z^0 \to q\bar{q}$	6.96	55%
Total Cross Section: 15.42 pb			

Table 5.2: MadGraph *cross section prediction for the processes* $p\bar{p} \to W\gamma$ *and* $p\bar{p} \to Z\gamma$. *The cross section for each boson channel decay is reported.*

5.3 Next to Leading Order Corrections

PYTHIA and MadGraph predictions are both based on LO matrix element calculations. However, theoretical studies [38] showed that the cross section of the $W(Z)\gamma$ process is enhanced when $O(\alpha_S)$ QCD contributions are included in the calculation. These Next-to-Leading Order (NLO) terms include both virtual subprocesses with gluon loops and real emission subprocesses with initial gluon or quark emissions (fig. 5.5). The correction to the cross section prediction is expressed as a factor $k = \frac{\sigma_{NLO}}{\sigma_{LO}}$ and its measurement was carried out in [37]. The σ_{NLO} was calculated with a NLO order $W\gamma$ and $Z\gamma$ Monte Carlo program [38] working in the narrow width approximation regime (*i.e.* not including the inner bremsstrahlung), a circumstance that matches very well our σ_{LO} definition.

The k-factor value is reported in fig. 5.6 as a function of P_T^γ and the resulting fit is:

$$\begin{cases} k = 1.62 + 0.0001 \cdot P_T^\gamma - 0.386 \cdot e^{-0.100 \cdot P_T^\gamma} & \texttt{for} \quad p\bar{p} \to W\gamma \\ k = 1.46 + 0.00073 \cdot P_T^\gamma - 0.125 \cdot e^{-0.062 \cdot P_T^\gamma} & \texttt{for} \quad p\bar{p} \to Z\gamma \end{cases}$$

By weighting the PYTHIA generated events with these k-factors and integrating over the whole P_T photon spectrum, we found an average NLO cross section correction of 1.55 for the $W\gamma$ process and 1.44 for the $Z\gamma$.

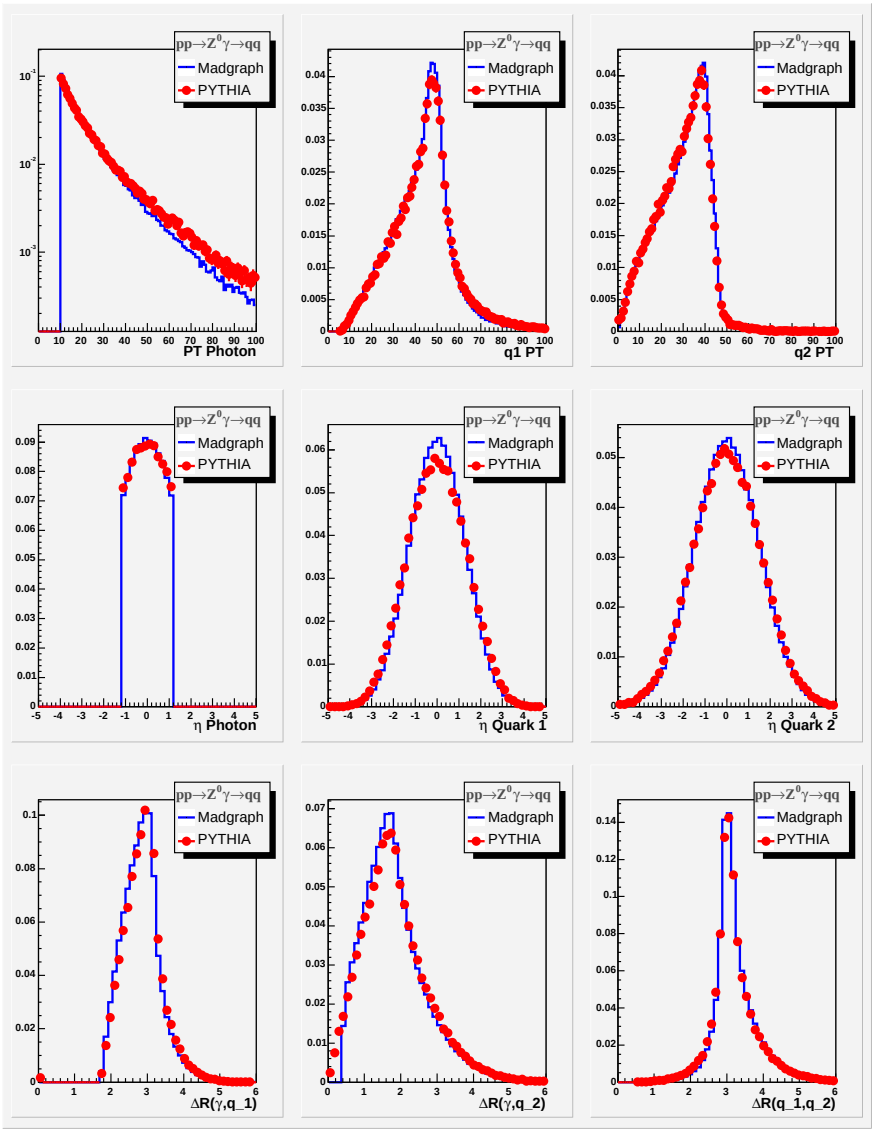


Figure 5.3: *PYTHIA*-MadGraph *comparison of three final parton kinematic distributions in* $p\bar{p} \to Z\gamma \to q\bar{q}\gamma$ *events. The momentum, pseudorapidity and the distance between each other* (ΔR) *are shown. On the* $\Delta R(\gamma, q_2)$ *distribution is visible the effect of the* MadGraph*-only generation cut* (>0.4).

5.4 Cross Section Systematic Errors

As described in the previous sections, the Monte Carlo program needs some input parameters to be set and some generation cuts to be imposed in order to

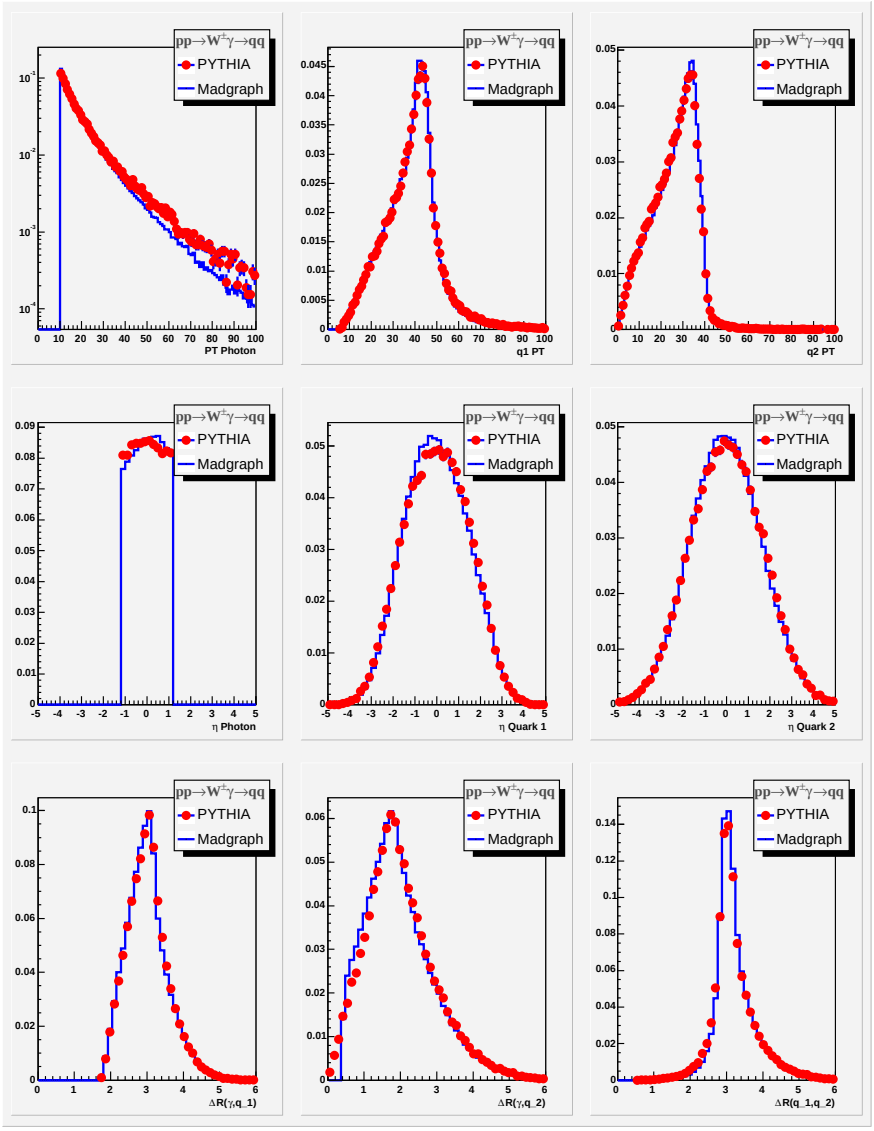


Figure 5.4: *PYTHIA*-MadGraph *comparison of three final parton kinematic distributions in* $p\bar{p} \to W\gamma \to q\bar{q}\gamma$ *events. The momentum, pseudorapidity and the distance between each other* (ΔR) *are shown. On the* $\Delta R(\gamma, q_2)$ *distribution is visible the effect of the* MadGraph*-only generation cut* (>0.4).

generate simulated events. The effects of these choices on the final cross section are part of the Monte Carlo prediction systematic uncertainties. They include also the k-factor errors and the PYTHIA-MadGraph difference. A discussion on

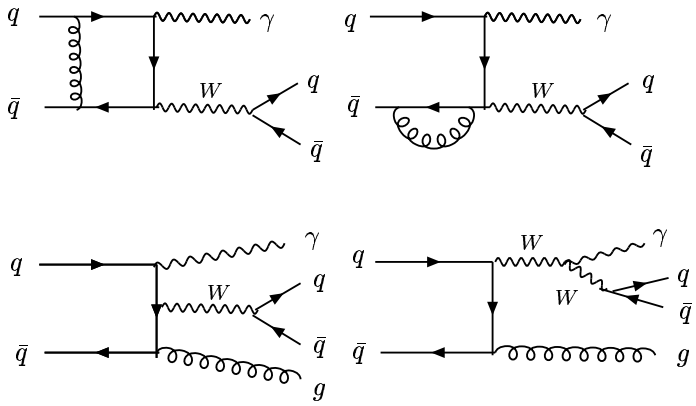
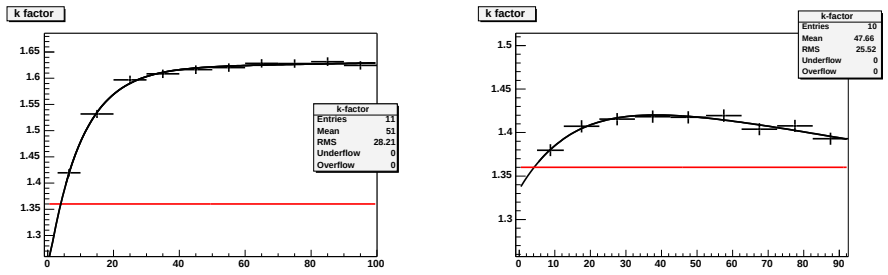


Figure 5.5: *Some examples of Feynman NLO diagrams contributions to $W(Z)\gamma$ production. The two diagrams on the top are example of virtual subprocess $q\bar{q} \to W\gamma \to q\bar{q}\gamma$ with initial state gluon loops. The two diagrams on the bottom are example of real emission subprocess $q\bar{q} \to W\gamma g \to q\bar{q}\gamma g$ with initial state gluon emission.*



(a) NLO $W\gamma$ correction

(b) NLO $Z\gamma$ correction

Figure 5.6: *Next to Leading Order (NLO) k-factor corrections for the $W\gamma$ and $Z\gamma$ production as reported in [37]. The average corrections are 1.55 and 1.44 for the $W\gamma$ and $Z\gamma$ respectively.*

each contribution to the systematics follows.

PDF: In both PYTHIA and MadGraph we used the CTEQ5L as LO parton distribution functions. Another common set of LO PDF are the MRST family [39]. To evaluate the impact of the PDF choice on the MC predictions we calculated the PYTHIA $W(Z)\gamma$ cross section using the MRST-72 PDF. We found a LO cross section $\sigma^{MRST-72} = 13.1$ pb, a 4.5% difference with respect to CTEQ5L. This variation is taken as the PDF choice systematic error.

Factorization Scale: The factorization scale q was set as the incoming quark momentum transfer $\sqrt{\hat{s}}$. Changing q by a factor 2 ($2\sqrt{\hat{s}}$, $\frac{1}{2}\sqrt{\hat{s}}$) results in a cross section variation of about 1.5%. We set this value as the LO cross section uncertainty due to the factorization scale.

Generation Cuts: Because of the infrared divergences at very low p_t photon we imposed P_T^γ >6 GeV in the Monte Carlo generation. It was the only kinematic cut we applied at the generation level in PYTHIA. To test if this cut affects the cross section prediction of $W(Z)\gamma$ events with P_T^γ >10 GeV, we produced a small sample with a 4< P_T^γ <6 GeV generation cut and we checked how many events were "promoted" above 10 GeV. Such amount of events turned out to be negligible (< 0.1%). Therefore we set to zero the systematic error associated to the generation cuts.

k-factor: The systematic error on the NLO corrections comes from the factorization scale, the statistical variation and the acceptance [3] In [40] the overall k-factor uncertainty was estimated to be 3% and since we use the same k-factor we set our systematic to this value as well.

Generator: Finally we associated a systematic error to the Monte Carlo generator choice and we set it equal to the difference in the cross section prediction between PYTHIA and MadGraph (11%). Since we used the the same PDF and factorization scale for both Monte Carlo and the generation cuts were very similar we assumed that this difference is due only to the way the two Monte Carlo manage the matrix element calculation. Therefore this source of systematic error is independent to the other ones and it can safely (no double counting) be added in quadrature to the previous contributions.

The cross section systematic uncertainties are summarized in Table. 5.3. The total uncertainty on the PYTHIA cross section prediction is 12.3%

[3]The presence of extra jets/gluons could affect the space distribution of the final partons

PYTHIA Cross Section Systematics	
PDF	4.5%
Factorization scale	1.5%
k-factor	3%
MC generator	11%
Total	12.3%

Table 5.3: *Summary of the systematic uncertainties of the PYTHIA* $q\bar{q} \to W(Z)\gamma \to q\bar{q}\gamma$ *MC cross section prediction. The total uncertainty is the sum in quadrature of all the contributions.*

5.5 Cross Section Summary

The PYTHIA Monte Carlo cross section prediction for the processes $p\bar{p} \to W\gamma \to q\bar{q}\gamma$ and $p\bar{p} \to Z\gamma \to q\bar{q}\gamma$ have been calculated in the region ($P_T^\gamma > 10$ GeV, $|\eta^\gamma| < 1.2$). Including the NLO corrections and the systematics errors it is:

$$\begin{aligned}\sigma_{W(Z)\gamma} &= \sigma(p\bar{p} \to W\gamma) \times \mathfrak{B}(W \to q\bar{q}) + \sigma(p\bar{p} \to Z\gamma) \times \mathfrak{B}(Z \to q\bar{q}) \\ &= 20.50 \pm 2.53 \;\; pb\end{aligned}$$

Chapter 6

Data Sample

In this chapter we will introduce the dataset [1] we used for our search of hadronic W/Z decay in photon events. In sec. 6.1 we will provide a short overview of the inclusive photon triggers implemented in CDF. We will also show that none of them has been proved useful for our analysis. In sec. 6.4 a new dedicated trigger, which we implemented in the CDF DAQ system, will be presented. The properties of this trigger and its efficiency in selecting the $W(Z)\gamma \to \gamma q\bar{q}$ signal events are reported in sec. 6.6.
The events selected by this trigger constitute the dataset on which we will perform our analysis.

6.1 Inclusive Photon Triggers

Specific triggers have been implemented in the CDF DAQ system to collect *inclusive* [2] photon events. They all follow a common strategy to select the photon: an EM cluster is required both at Level 1 (sec. 4.3) and at Level 2 (sec. 4.4.1) and some quality cuts are applied to the Level-3 reconstructed photon. What changes among these triggers is the photon E_T threshold and the isolation cuts applied at the different stages. Because the inclusive photon cross section diverges when the photon E_T approaches zero, the rate of the photon trigger data has to be reduced by either setting a high E_T threshold or collecting data with a suitable prescale factor [3]. Both these options have been adopted resulting in 4 inclusive photon triggers with E_T threshold of 8, 10, 15

[1] A dataset is a sample containing all the events that passed some particular trigger. Sometime, for the sake of brevity, we will just use the name of the trigger instead of "the dataset corresponding to...".

[2] For inclusive we mean triggers where only the presence of at least one photon is required. Also not inclusive photon triggers are implemented requiring - besides the photon - the presence of other objects like for instance another photon, a muon, a b-jet, etc...

[3] When a prescale factor is set to n, only 1 every n events passing the trigger is written on tape. If all the events passing the trigger are actually recorded, the trigger is called $unprescaled$ ($n = 1$).

and 25 GeV. Different prescale factors are applied to them and currently only the last one is unprescaled.
The specifics of those triggers are summarized in Table 6.1.

Inclusive Photon Triggers				
Trigger Name	PHOTON_EM8	PHOTON_10_ISO	PHOTON_15_TIGHT	PHOTON_25_ISO
Level 1				
E_T (GeV)	> 8	> 8	> 8	> 8
E_{HAD}/E_{EM}	< 0.125	< 0.125	< 0.125	< 0.125
Level 2				
E_T (GeV)	-	> 10	> 12	> 21
E_{HAD}/E_{EM}	-	< 0.125	< 0.125	< 0.125
`bestiso` (GeV)	-	$\min(0.15 \cdot E_T, 3)$	< 1.0	< 3.0
Region	-	-	Central	-
Level 3				
E_T (GeV)	-	-	> 15	> 25
E_{HAD}/E_{EM}	-	-	< 0.05	$(0.055+$ $0.00045 \cdot E_T)$
E_T^{Iso} (GeV)	-	-	< 1.0	< 2.0
χ^2_{CES}	-	-	< 20	< 20
$\lvert X_{CES}\rvert$ (cm)	-	-	< 21	-
$\lvert Z_{CES}\rvert$ (cm)	-	-	$9 < z < 217$	-
Prescale	2500	1600	50:1	1

Table 6.1: *Inclusive photon trigger specifics. The Level-1 and Level-2 quantity reconstruction are described in sec. 4.3 and sec. 4.4.1 respectively. The Level-3 quantities are reconstructed with the same algorithms used offline and described in sec. 4.7.1. The* PHOTON_15_TIGHT *trigger is* dynamically *prescaled (see sec. 6.7).*

6.2 The Photon plus Two Jets Trigger

As briefly mentioned previously, besides the inclusive photon triggers described above several other *non-inclusive* photon triggers are implemented. One of them - PHOTON_18_DIJET trigger - requires at Level3 two jets with an energy of 18 GeV and 10 GeV respectively. As far as the photon selection is concerned it implements the same cuts of the PHOTON_25_ISO trigger but with a Level-3 E_T threshold of 18 GeV instead of 25 GeV and, most importantly, it is unprescaled. Because of its jet requirements this trigger seems a good candidate for the search of the $W(Z)\gamma \to \gamma q\bar{q}$ events.

6.3 Choice of the trigger

In order to decide which is the best dataset to use for our analysis (*i.e.* the dataset that is likely to contain the highest number of signal events) we need to know as the signal cross section depends on the P_T^γ. Using the MadGraph Monte Carlo sample introduced in sec. 5.2.2 we plot the relative cross section - normalized at 25 GeV - as a function of the P_T^γ (fig. 6.1).

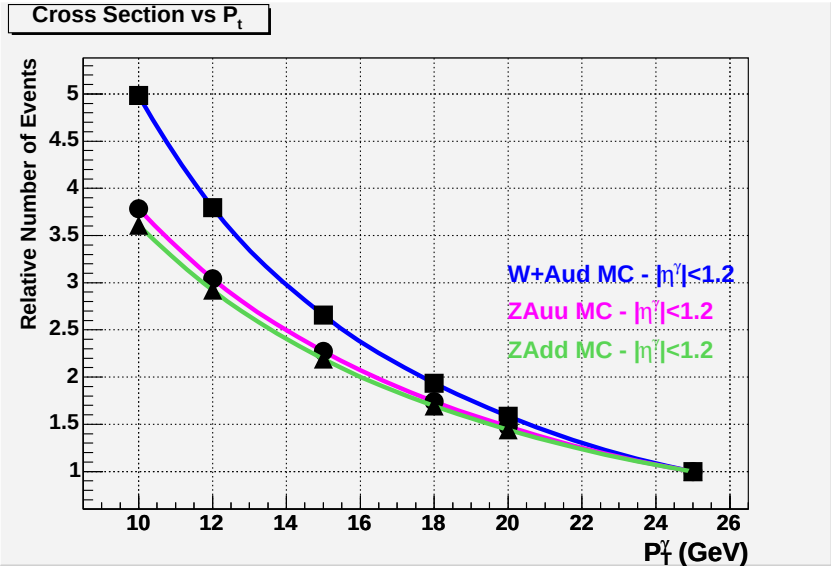


Figure 6.1: *Relative cross section of* $p\bar{p} \to W(Z)\gamma \to \gamma q\bar{q}$ *process as resulting from the* MadGraph *Monte Carlo calculation. Only the channels* $W^+\gamma \to u\bar{d}$ (■), $Z^0\gamma \to u\bar{u}$ (●) *and* $Z^0\gamma \to d\bar{d}$ (▲) *are shown. The photon is required to have* $|\eta^\gamma| < 1.2$.

As the inclusive photon production, the $W(Z)\gamma \to \gamma q\bar{q}$ cross section exponentially increases at low P_T^γ, being at $P_T^\gamma = 10$ GeV about five times higher than at $P_T^\gamma = 25$ GeV. Accounting for this behavior, the dataset with the lowest threshold seems to be the best choice. Unfortunately big prescale factors heavily suppress the amount of collected events and this loss is not compensated by the enhanced signal yield.
Ruling out the prescaled triggers the possible choices left are the `PHOTON_25_ISO` and the `PHOTON_18_DIJET` triggers. The latter has a lower photon threshold but it is also not inclusive (it requires two jets at Level 3). Therefore is not granted *a priori* that its signal yield would be larger than the `PHOTON_25_ISO` one, in spite of the higher threshold of the latter.
But there is another criteria we must require to a dataset in order to be suitable for the $W(Z)\gamma \to \gamma q\bar{q}$ search. As we have emphasized in Ch.5 to derive the background shape of the dijetmass distribution in the signal region (in our case

between 60 and 120 GeV) we have to fit both the left and the right sidebands, [4] and extrapolate them. This can be successfully done only if the turn-on of the dijetmass distribution is low enough *to allow* the presence of the left sideband. As can be seen in fig. 6.2 the resulting turn on of the dijetmass distribution for both the `PHOTON_25_ISO` and `PHOTON_18_DIJET` dataset is s close to the 60 GeV limit. This imply that *both* dataset are not appropriate for the $W(Z)\gamma \to \gamma q\bar{q}$ search (at least in the way we plan to do it).
In conclusion, none of the implemented photon triggers can be used for our analysis leaving as only alternative the need to device a new dedicated trigger able to effectively collect a large sample of $W(Z)\gamma \to \gamma q\bar{q}$ events with a reasonable trigger rate.
In the following sections we will describe in details such a new trigger we implemented in the CDF DAQ system and optimized for the selection of $W(Z)\gamma \to \gamma q\bar{q}$ events.

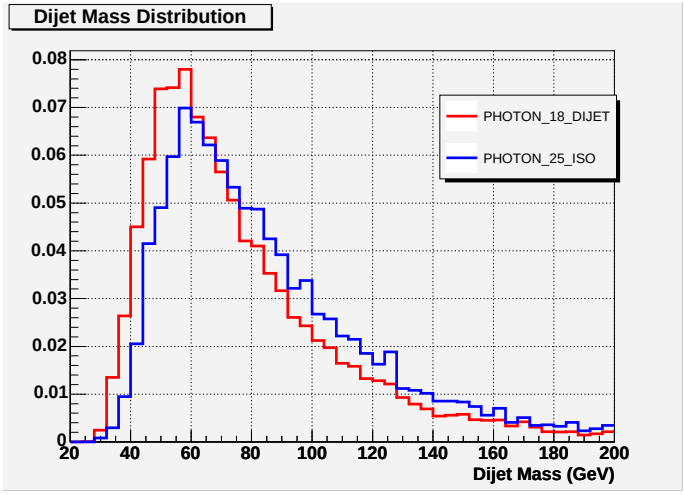


Figure 6.2: *Dijetmass distribution of events taken with the* `PHOTON_18_DIJET` *trigger (blue) and* `PHOTON_25_ISO` *(red). The* $W(Z)\gamma \to \gamma q\bar{q}$ *signal region starts at 60 GeV.*

6.4 A new photon trigger

As outlined in the previous section the main prerequisites a new trigger need to have in order to be effective for our analysis are:

❶ High Signal Yield: It has to collect a large amount of $W(Z)\gamma \to \gamma q\bar{q}$ events. Therefore it must have a low P_T^{γ} threshold and to be unprescaled (or prescaled with a small factor).

[4] the part of the distribution outside the interested region where only the background contributes.

❷ Low Rate: It must have a reasonable trigger rate (roughly below 100 nb [5] at Level 2 and below 25 nb at Level 3)

❸ Low Turn-On: The dijetmass distribution turn-on must be lower than the signal region limit of 60 GeV to allow a background fit of the left sideband.

It has to be noticed that both ❶ and ❸ favor a low P_T^γ threshold. Unfortunately this is in striking contrast with ❷, since the rate grows exponentially at low P_T^γ [6]. It is clear that to keep the rate under control *and* still have a low P_T^γ threshold it is mandatory to heavily filter the events at Level-2 [7] introducing further requirements to enhance the purity of the sample.
The first thing to be introduced is the requirement for the photon to fall in the central calorimeter (in this way we can also employ the photon background subtraction method described in sec. 4.7.1). Together with a tight isolation cut of 1 GeV - as in the `PHOTON_15_TIGHT` trigger - these two cuts provide a substantial reduction in rate with a very good photon selection efficiency. Because no much more can be done on the photon selection, in order to further reduce the trigger rate we investigated the possibility of exploit the presence - in the signal events - of two jets coming from the decays of a heavy boson. This will be discussed next.

6.4.1 Jet Requirement at Level 2

Since in the signal events we expect at least two jets with relatively high energy, it seems convenient to require two L2 jets at the trigger level. But as briefly mentioned in sec. 4.4.2 because of the limited information available at this stage and the large uncertainty associated with them, the L2 jet reconstruction is highly inefficient (fig. 6.3). Moreover even when it does find a jet, the associated energy can substantially differs from the energy reconstructed at the offline level (fig. 6.4). This is not too surprising considering the kind of clustering algorithm employed at Level 2.
Nevertheless it is plausible that at least the most energetic jet from the W(Z) decay (typically produced with an average energy of 40 GeV) is still reconstructed at Level 2. For this reason we added to the trigger requirements the presence of at least one L2 jet but *without* any energy threshold (because we cannot rely on the Level-2 measurement of the jet energy). Furthermore since the two jets from the W(Z) in the signal events are produced mainly in the central region [8] (with a falling distribution for $|\eta|$ >1.5) we can also apply an η cut on the same L2 jet. On the other hand looking at the L2 jet η distribution

[5] It is customary to assess the trigger rate in barn. The correspondent rate (in Herz) is $N(Hz) = \sigma(nb) \cdot L(10^{31})/100$ if σ is expressed in nb and the luminosity L in 10^{31} cm^{-1} s^{-1} units.

[6] To give an example the `PHOTON_10_ISO` trigger has an impressive Level-2 rate (before prescale) of 3600nb!

[7] The rate bottleneck is actually only at Level-2. At Level-1 the `PHOTON_EM_8` is unprescaled and at Level-3 a filtering close to the offline analysis cuts can be introduced if needed.

[8] This is a consequence of the central photon selection and of the high mass of the W with respect to the p_T of the photon.

in the inclusive PHOTON_10_ISO dataset (fig, 6.5) we see an enhanced activity at high eta (caused by the forward QCD jets) which very unlikely has something to do with the W(Z) production. We therefore implemented a $|\eta^{L2jet}|$ <1.78 requirement. It has to be noticed that the signal efficiency selection is enhanced by the presence of two jets in the in the final state of the signal events because the probability that at least one of them is reconstructed at L2 is twice as much that in a single jet case.

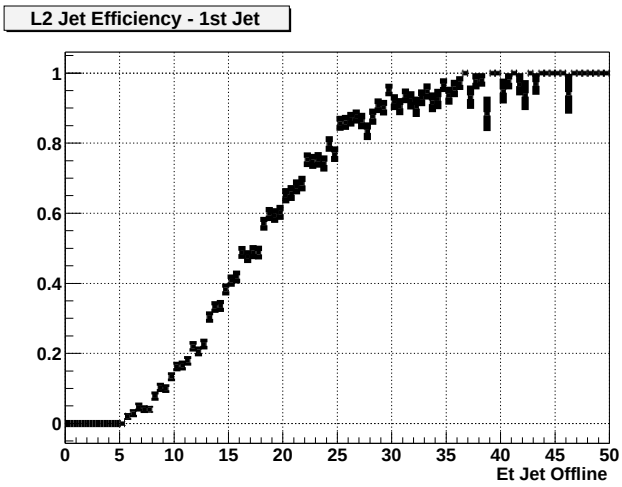


Figure 6.3: *The L2 jet efficiency as a function of the offline jet E_T (corrected). We defined the efficiency as the fraction of offline jets that match ($\Delta R(jet^{OFF}, jet^{L2}) <$ 0.4) a L2 jet (of any energy).*

6.4.2 Sumet requirement at Level 2

Another quantity that can be effectively used to reduce the rate at Level 2 is the L1 sumet [9] (see sec. 4.3). The inaccuracy of the L2 jet energy measurement didn't allow us to take advantage of the high E_T value of the W(Z) jets. On the other hand this implies also an higher sumet for a signal event with respect to a typical photon event (see fig. 6.6). The L1 sumet as well suffers its own uncertainties and thus it is recommended to implement only a loose cut on it. Even a better quantity to cut on is the total E_T of the event *excluding* the contribution of the photon. It is more sensitive to the presence of energetic jets and more effective in reducing the rate. Therefore a $\sum_{L1} E_T - P_T^{\gamma} > 20$ GeV cut was implemented in the new trigger.
Although we investigated other L2 information, the final version of the trigger

[9] This quantity is actually calculated at L1 and can be in principle be used at that level. But for practical reason it is better to use an already existent L1 unprescaled trigger and use the L1 sumet information at L2. The filtering power remains exactly the same.

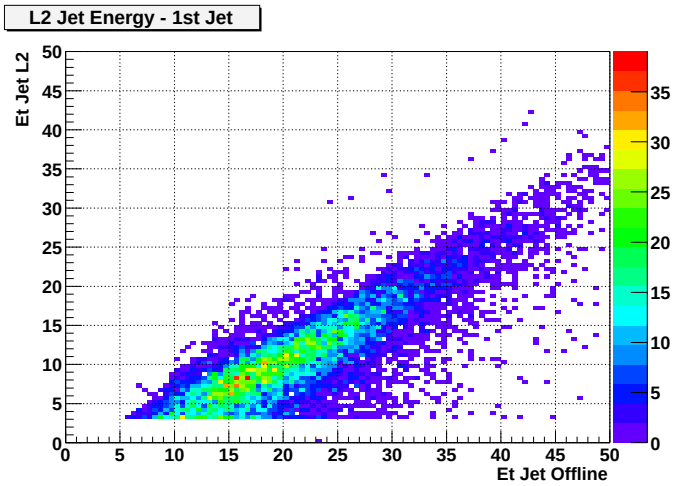


Figure 6.4: *Scatter plot of the offline jet energy (corrected) versus the L2 jet energy, when the two jets match ($\Delta R(jet^{OFF}, jet^{L2}) < 0.4$). It is clear that for a L2 jet of a given E_T the corresponding offline E_T can differs substantially.*

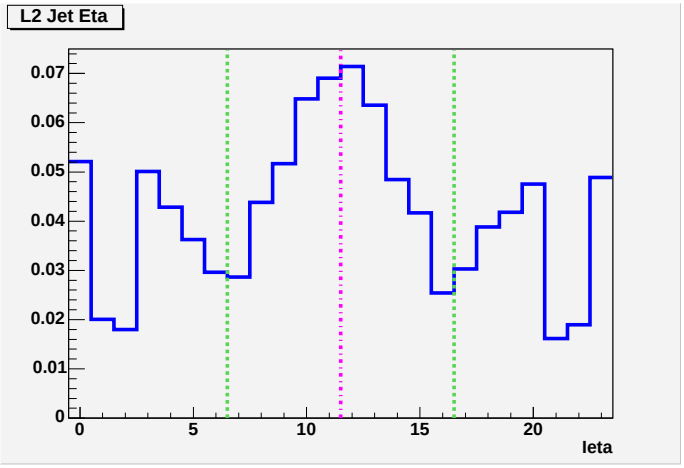


Figure 6.5: *L2 Jet η distribution for photon inclusive events. What is plotted is actually* `ieta`*, the trigger tower index. The* $0 \leq$ `ieta` ≤ 11 *region corresponds to the west side of the detector, The central calorimeter corresponds to* $7 \leq$ `ieta` ≤ 16*. The region* $|\eta| < 1.78$ *corresponds to* $4 \leq$ `ieta` ≤ 20*. An enhanced activity is observed in the forward region.*

has the requirements described above and summarized in Table 6.2 (the reason of the two versions will be explained in the following section).

Finally we need to check if all the extra requirements we added to the trigger affected the position of the dijetmass turn on. In fig. 6.7 we compared the dijetmass distribution of events passing the new trigger and the `PHOTON_18_DIJET` trigger. It is clear that we succeed in conceiving a trigger well suited for the $W(Z)\gamma \to \gamma q\bar{q}$ search analysis.

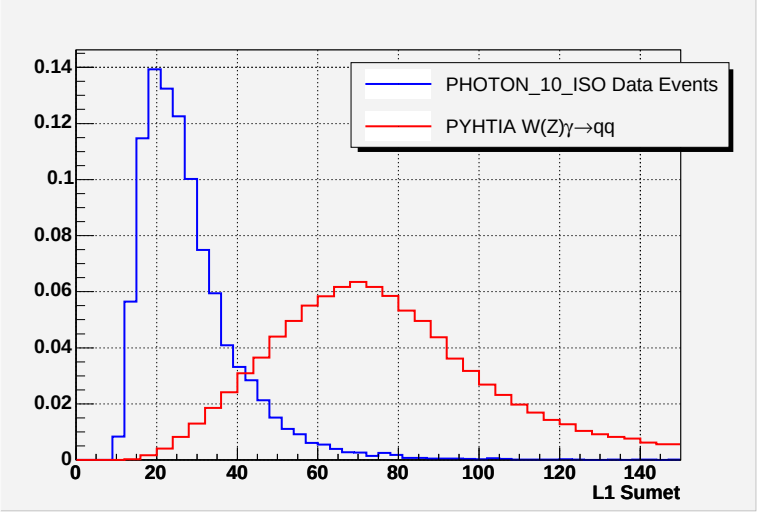


Figure 6.6: *L1 sumet distribution for data event taken with the inclusive* `PHOTON_10_ISO` *trigger (blue) compared to the L1 sumet of the* $W(Z)\gamma \to \gamma q\bar{q}$ *signal events.*

6.5 Implementation History of the New Trigger

A first version of the new trigger was officially included in the CDF DAQ system at the end of July 2003. It replaced the `PHOTON_18_DIJET` trigger - being an improved version of it - with the new name "`PHOTON_DIJET_L2_DPS` " and the photon threshold was set to 10 GeV. As its name suggests, it was dynamically prescaled (DPS). For this particular trigger it implied to be unprescaled except for instantaneous luminosity higher than $5 \cdot 10^{30}$ $cm^{-2}s^{-1}$. In this case a prescale factor of 5 were applied. As we will show in sec. 6.7 this did not cause a big event loss. The decision to apply a DPS was driven by the high L2 trigger rate (with respect to the bandwidth budget assigned to it) and by the fact that it caused some troubles only at high luminosity (that, at the time this version of the trigger was operating, was seldom reached). However the increasing performance of the accelerator prompted the decision, at the end of January 2004, of removing the DPS and compensate the higher rate increasing the photon threshold up to 12 GeV. At this second version of the trigger was assigned the name "`PHOTON_DIJET` ".

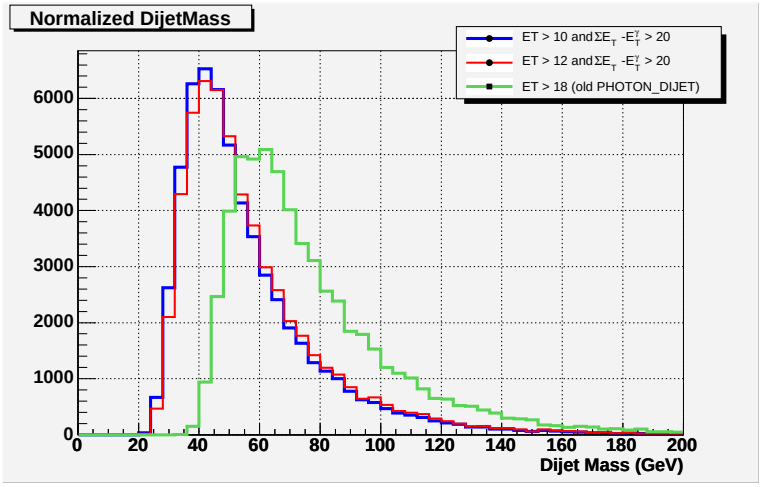


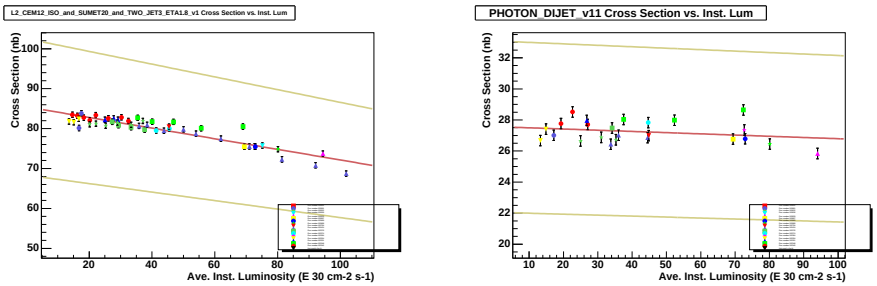
Figure 6.7: *The dijetmass distribution of events taken with the new trigger when the P_T^γ threshold has been set to 10 GeV (blue) and 12 GeV (red). The turn-on is clearly well below the start of the signal region at 60 GeV. As a comparison the dijetmass distribution of the previous* `PHOTON_18_DIJET` *trigger is shown. In this case the turn-on is inside the signal region, making this trigger not suitable for the $W(Z)\gamma \to \gamma q\bar{q}$ search analysis.*

In Table 6.2 we summarized the two version trigger definitions and in Table 6.3 we outlined their implementation history along with the total integrated luminosity collected by the two samples. Both these datasets are employed for the search of the $W(Z)\gamma \to \gamma q\bar{q}$ events reported in this dissertation. The `PHOTON_DIJET` trigger rate at Level 2 and at Level 3 as a function of the instantaneous luminosity is reported in fig. 6.8. From these plots results that the trigger rate *does not* increase at high luminosity and at Level 3 it remains constant for all luminosity values [10].

6.6 Trigger Efficiency

After having checked that the trigger has an acceptable rate, we have now to see how much it is efficient in selecting our $W(Z)\gamma \to \gamma q\bar{q}$ signal. Although we were very careful to include L2 cuts with a minimal impact on signal events, a fraction of them will not pass the requirements we implemented in the new trigger. To measure this fraction (the "*trigger efficiency*") we used both data and Montecarlo samples.
Since this trigger selects photons and jets we can separate the two contributions to the total efficiency ϵ_{trg}:

[10] This is due to the tight isolation cuts of this trigger that tends to reject events with a large amount of pile-up energy.



(a) Level-2 Rate

(b) Level-3 Rate

Figure 6.8: *The rate (expressed in nb) of the* `PHOTON_DIJET` *at Level 2 (a) and at Level 3 (b) as a function of the instantaneous luminosity. It can be noticed that the rate of this trigger is basically constant with the luminosity at Level 3 and slightly decreases at Level 2. The colors and points refer to different run number intervals*

- ϵ_{ph}, defined as the fraction of offline photons with $P_T^\gamma > 12$ GeV (see sec. 4.7.1) passing the photon part of the trigger requirements (the L1 and L3 cuts plus the L2 photon cuts as reported in Table 6.2);
- ϵ_{jet}, defined as the fraction of $W(Z)\gamma \to \gamma q\bar{q}$ signal events passing the L2 jet requirements (the L2 jet cuts on Table 6.2).

This separation is particularly convenient because the selection of the photon in our signal sample is not different from the typical prompt photon selection of other analyses [40] and thus it is very well understood. On the other hand the jet selection efficiency is peculiar of this analysis only (the jets in our signal are coming from the decay of heavy bosons).
Because the two versions of the trigger have some differences (the L2 photon threshold and the $\sum_{L1} E_T$) we need to calculate their efficiencies separately.

6.6.1 Photon Trigger Efficiency (ϵ_{ph})

We have defined the probability ϵ_{ph} for an offline photon to pass the photon subset of the trigger requirements. They differ from those ones implemented in the `PHOTON_10_ISO` trigger for the L2 isolation cut (`L2_iso`), the L3 cuts (`L3_TIGHT`) and - for the `PHOTON_DIJET` only - the L2 photon threshold (`L2_clus`). As briefly mentioned previously, the `PHOTON_10_ISO` trigger efficiency ϵ_{10iso} has been already measured in other studies [11]. Therefore we can derive our ϵ_{ph} using the the `PHOTON_10_ISO` dataset and calculating the fraction of events passing the tighter photon cuts of the `PHOTON_DIJET_L2_DPS` and `PHOTON_DIJET` triggers.

[11] The quoted number is $\epsilon_{10iso} = a \cdot \texttt{Erfc}(b \cdot (c - P_T^\gamma))$ with a=0.0498, b=0.0649 and c=12.31. The function `Erfc` is defined $Erfc(x) = \frac{1}{2\sqrt{\pi}} \cdot \int_x^\infty e^{\frac{t^2}{2}} dt$

The New Photon Plus Jets Trigger		
Trigger Name	PHOTON_DIJET_L2_DPS	PHOTON_DIJET
Level 1		
E_T (GeV)	> 8	> 8
E_{HAD}/E_{EM}	< 0.125	< 0.125
Level 2 - Photon Cuts		
E_T (GeV)	> 10	> 12
E_{HAD}/E_{EM}	< 0.125	< 0.125
bestiso (GeV)	< 1.0	< 1.0
Region	Central	Central
Level 2 - Jet Cuts		
$\sum_{l1} E_T$ (GeV)	> 30	$> 20 + P_T^{\gamma}$
L2 Jet	> 1	> 1
L2 Jet $\lvert\eta\rvert$	< 1.78	< 1.78
Level 3		
E_T (GeV)	> 10	> 10
E_{HAD}/E_{EM}	< 0.05	< 0.05
E_T^{Iso} (GeV)	< 1.0	< 1.0
χ^2_{CES}	< 20	< 20
$\lvert X_{CES}\rvert$ (cm)	< 21	< 21
$\lvert Z_{CES}\rvert$ (cm)	$9 < z < 217$	$9 < z < 217$
Prescale	5:1	1

Table 6.2: *Specifics of the two versions of the new photon plus jets triggers optimized for the search of* $W(Z)\gamma \to \gamma q\bar{q}$ *events. In the second version the photon* P_T *threshold was raised to 12 GeV at Level-2 to reduce the trigger rate and allow the removal of the DPS. Also the* $\sum E_T$ *cut was changed to further reduce the rate.*

We call these efficiencies $\epsilon_{dj10\to 10iso}$ and $\epsilon_{dj12\to 10iso}$ respectively. For the total photon efficiency we will simply have

$$\begin{cases} \epsilon_{ph} = \epsilon_{dj10\to 10iso} \cdot \epsilon_{10iso} & \text{for the PHOTON\_DIJET\_L2\_DPS} \\ \epsilon_{ph} = \epsilon_{dj12\to 10iso} \cdot \epsilon_{10iso} & \text{for the PHOTON\_DIJET} \end{cases}$$

As far as the L2 cuts are concerned (L2_iso and L2_clus), since the information of the L2 electromagnetic cluster is available offline, we know *exactly* (no systematic errors) how many events satisfy L2_iso< 1 GeV and L2_clus>12 GeV. We reported the efficiency of these cuts in the two plots on the top of fig. 6.9 and fig. 6.10. For the L3 cuts instead we applied them directly to the corresponding offline quantities. This is justified by the fact that - as we explained in sec. 4.5 - the L3 and offline objects use the same reconstruction algorithms. However we know that the detector information (calorimeter energy, vertex position, etc...)

New Trigger Implementation History			
Period	*Run Range*	*Integrated Luminosity* (pb^{-1})	
		PHOTON_DIJET_L2_DPS	PHOTON_DIJET
Aug. 2003 - Sep. 2003	166677-168889	29.5	-
Nov. 2003 - Jan. 2004	175066-178744	24.4	-
Jan. 2004 - Feb. 2004	178683-179056	-	18.4
May 2004 - Aug. 2004	182843-186598	-	111.7
Total Integrated Luminosity		**184 pb^{-1}**	
Total Number of Events			
Fraction of events passing PHOTON_10_ISO		3%	0.06%
Fraction of events passing PHOTON_15_TIGHT		17.1%	18.4%
Fraction of events passing PHOTON_25_ISO		0.9%	2.4%

Table 6.3: *New photon plus jets trigger history. The integrated luminosity collected by the* PHOTON_DIJET_L2_DPS *trigger is* corrected *by its effective prescale (sec. 6.7). The fraction of events collected by the new triggers and passing also other inclusive photon triggers is reported.*

are not exactly the same and therefore we assigned a 2% systematics at this part of the efficiency measurement.
The efficiency for the L3 cuts is reported in fig. 6.9 (bottom left plot) for the PHOTON_DIJET trigger along with the combined efficiency $\epsilon_{dj12\to 10iso}(P_T^\gamma)$. The correspondent plots for the PHOTON_DIJET_L2_DPS trigger are in fig. 6.10. In this case the $\epsilon_{dj10\to 10iso}(P_T^\gamma)$ does not depend upon the P_T^γ because of its lower threshold.

6.6.2 Jet Trigger Efficiency (ϵ_{jet})

To evaluate the ϵ_{jet} efficiency we need to gauge the effect of the Level-2 jet cuts on the $W(Z)\gamma \to \gamma q\bar{q}$ signal events. Since only Monte Carlo samples are available for these events, the Level-2 object information are not available, as they are related only to the data taking process [12]. To by-pass this problem we tried to emulate the trigger object reconstruction processes using only the offline information. First of all we noticed that we use - in the trigger selection - only the calorimeter information and they are not very different from the offline counterpart except for some small correction factors applied during the offline calibration and for the different geometry of the towers. At the trigger level indeed the calorimeter tower grid is simplified from an offline 52x24 for the central and 52x48 for the plug (see sec 3.3) to a evenly 24x24 grid obtained gathering nearby towers.

[12] Actually a trigger simulation software has been developed to generate the output of the trigger system in Monte Carlo events but we decided to not use it to avoid complications in the systematics calculation.

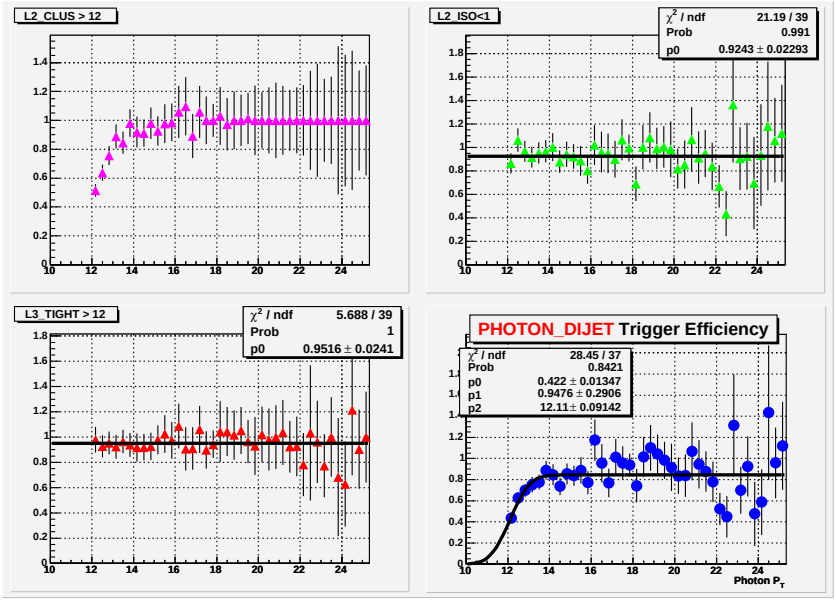


Figure 6.9: `PHOTON_DIJET` *trigger efficiencies of the L2 photon cuts* `L2_clus`$>$*12 GeV and* `L2_iso`$<$ *1 GeV (top plots) and of the L3 cuts (bottom left plot). The dataset used is the* `PHOTON_10_ISO` *trigger. The bottom right plot is the* $\epsilon_{dj12\to 10iso}(P_T^\gamma)$*. The efficiency is parametrized with the function* $\epsilon_{dj12\to 10iso}(P_T^\gamma) = p_0 \cdot \mathtt{Erfc}(p_1 \cdot (p_2 - P_T^\gamma))$ *and the result of the fit is reported in the plot inset.*

Using the offline tower energies and applying the same tower rearrangement we were able to "*emulate*" the calorimeter information as they are at the trigger level. To check how well we reproduced the trigger information, we applied our method to the data and we compared the real tower energies with the"*emulated*"ones. As shown in fig. 6.11(a) the agreement is superb with a resolution of about 75 MeV per tower.
However since we actually use the L2 Jet and the $\sum_{L1} E_T$ information we should verify how well we can reproduce them with our emulation procedure. Unfortunately it is not practically feasible - although technically possible - to emulate the complex L2 jet reconstruction algorithm (see sec. 4.4.2). But, thanks to our particular cuts, it is not even necessary. Indeed in the new photon plus jet triggers we just require the presence of a L2 jet of any energy and with an $|\eta|$ <1.78. In particular this means that an event with at least one trigger tower with $E_T > E_T^{seed}$ in the region $|\eta|$ <1.78 would pass the L2 jet trigger cuts. Expressed in this way, the L2 jet cuts can be easily applied to the "*emulated*" towers. As a further check of the correctness of our argument we plot the distribution (fig. 6.11(b))of the most energetic "*emulated*" trigger tower for data events *without* any reconstructed L2 jet. As expected, for these events there is basically no "*emulated*" towers with an energy above $E_T^{seed} = 3$ GeV.

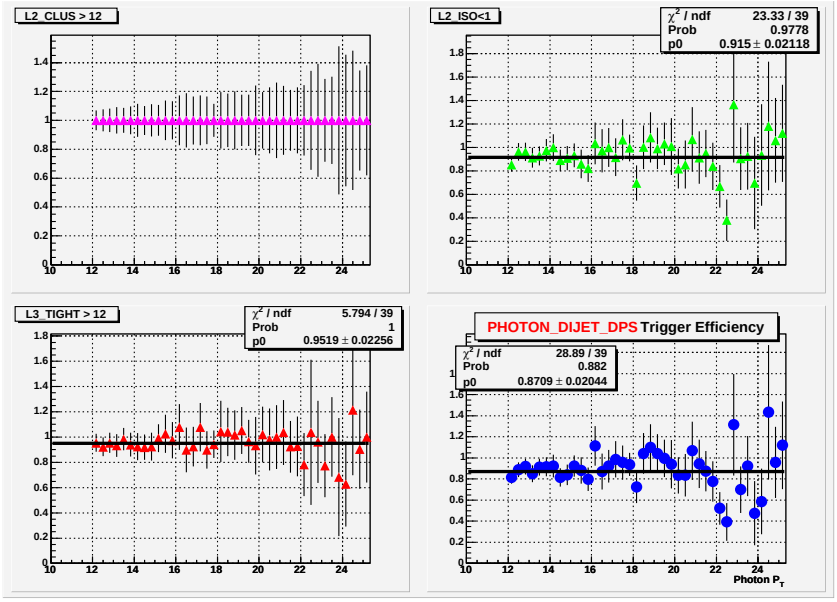


Figure 6.10: `PHOTON_DIJET_L2_DPS` *trigger efficiencies of the L2 photon cuts* `L2_iso`$<$ *1 GeV (top plots) (the* `L2_clus`$>$*12 GeV is 100% efficient because this trigger has the same L2* E_T *threshold of the* `PHOTON_10_ISO` *trigger) and of the L3 cuts (bottom left plot). The dataset used is the* `PHOTON_10_ISO` *trigger. The bottom right plot is the* $\epsilon_{dj10\to 10iso}(P_T^\gamma)$.

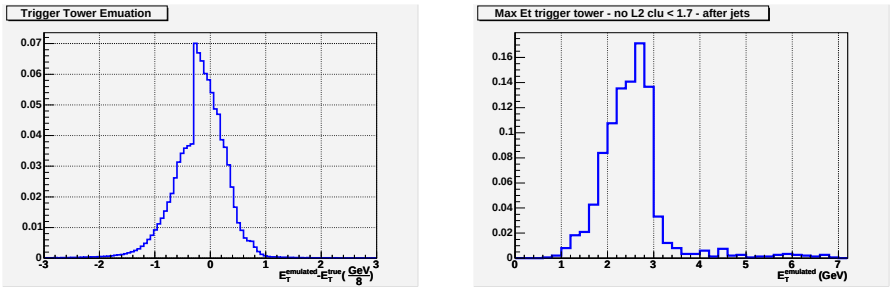
To check if we could reproduce the quantity $\sum_{L1} E_T$ as well, we then applied exactly the same trigger algorithm [13] to the "*emulated*" trigger towers. With no surprise we found an excellent agreement when we compared the "*emulated*" $\sum_{L1} E_T$ value with the real one as calculated by the trigger system (fig. 6.12). A similar agreement is also achieved for the quantity used in the `PHOTON_DIJET` trigger, $\sum_{L1} E_T - P_T^\gamma$. Using different dataset we estimated the "*emulated*" $\sum_{L1} E_T$ precision to be within 1.5 GeV.

Now that we are confident that we can reliably reconstruct the trigger cuts using the offline calorimeter information, we can look at their distributions in the Monte Carlo signal events and finally assess the jet part of the trigger efficiency ϵ_{jet}. Assuming the L2 jet and the $\sum_{L1} E_T$ cuts independent we can break up the efficiency and write:

$$\epsilon_{jet} = \epsilon_{sumet} \cdot \epsilon_{seed}$$

The ϵ_{seed} is the efficiency of the L2 jet requirements that - as discussed before - are equivalent to the presence of at least one trigger tower above the

[13] The $\sum_{L1} E_T$ is not simply the scalar sum over all the trigger tower of their transverse energy. Because the algorithm is encoded in the hardware some weird simplifications (like the drop of the last two digits of the energy bit) are applied making impossible to reproduce this number unless exactly the same algorithm is applied.



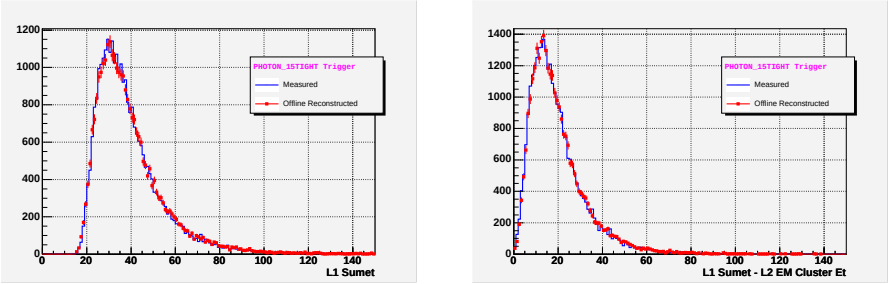
(a) (b)

Figure 6.11: *(a) Difference between the "*emulated*" energy and the measured energy of trigger towers for data events. One unit here is 125 MeV. The strange asymmetry is due to the fact that the calorimeter energy at the trigger level is recorded in 125 MeV unit. (b) Distribution of the "*emulated*" energy for the most energetic tower in the $|\eta|$ <1.78 region for data events that donot pass the L2 jets requirements. This plot shows that in this case there are not basically any "*emulated*" towers above 3 GeV, an equivalent condition for a L2 jet veto.*

seed energy (3 GeV) and located in the region $|\eta|$ <1.78. To evaluate ϵ_{seed} we plot the E_T distribution of the most energetic "*emulated*" tower (E_T^{max}) with $|\eta|$ <1.78 for the signal $W(Z)\gamma \rightarrow \gamma q\bar{q}$ events. Then, simply calculating the fraction of events with $E_T^{max} > 3$ GeV we get ϵ_{seed}. Using the plot in fig. 6.13 we found an efficiency of 98.7%. Since the "*emulation*" method gives the corrected energy of a trigger tower within 125 MeV we consider the systematic associated with this uncertainty negligible.
For the ϵ_{sumet} we need to calculate two different efficiencies - ϵ_{sumet}^{DPS} and ϵ_{sumet}^{dj12} - for the two versions of the triggers. In the `PHOTON_DIJET_L2_DPS` case the trigger cut is $\sum_{L1} E_T$ >30 GeV. Plotting the corresponding "*emulated*" distribution for signal events (fig. 6.14(a)) we can calculate the fraction of events above the 30 GeV threshold and set ϵ_{sumet}^{DPS}. The value we found is 98.0%. We repeated the same procedure for the quantity $\sum_{L1} E_T - P_T^{\gamma}$ with a 20 GeV threshold (fig. 6.14(b)). In this case we found $\epsilon_{sumet}^{dj12} = 94.6\%$.
Shifting both threshold of 1.5 GeV (the accuracy we quoted for the "*emulated*" $\sum_{L1} E_T$) we found the two efficiencies not moving more than 1% that we set as systematic uncertanty.

6.7 The Dynamically Prescale Factor

The last thing we need to do to complete the new trigger efficiency study is to measure the "*effective*" prescale applied to the `PHOTON_DIJET_L2_DPS` trigger. For most of the time it was part of the data acquisition system, this trigger was set in an unprescaled mode. However when the experiment was running at



(a) $\sum_{L1} E_T$ (b) $\sum_{L1} E_T - P_T^{\gamma}$

Figure 6.12: *(a) This plot shows the* $\sum_{L1} E_T$ *distribution measured at the trigger level (blue curve) and the same quantity reconstructed using the "*emulated*" trigger tower energy (red circle) for* PHOTON_15_TIGHT *trigger events. (b) Same comparison for the* $\sum_{L1} E_T - P_T^{\gamma}$ *quantity.*

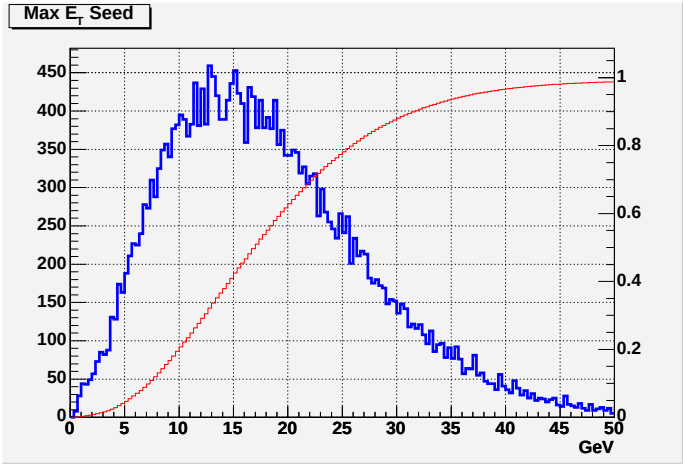
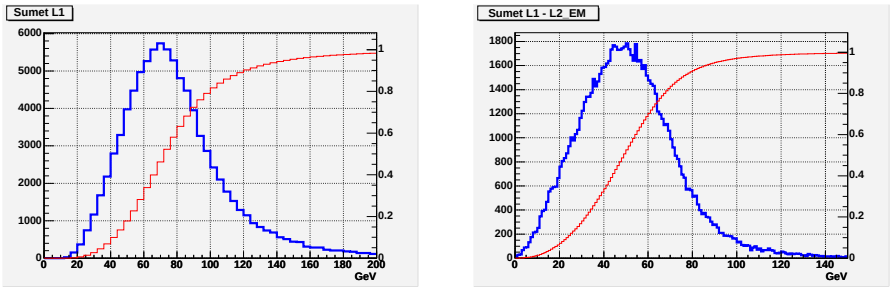


Figure 6.13: E_T^{max} *distribution of highest "*emulated*" energetic tower in the* $|\eta|$ *<1.78 region for* $W(Z)\gamma \rightarrow \gamma q\bar{q}$ *Monte Carlo events. The threshold for the L2 jet seed tower is 3 GeV. The events with* E_T^{max} *>3 GeV would pass the L2 jet requirements implemented in the photon plus jet triggers.*

high luminosity a prescale factor 5 was applied in order to save precious Level 2 bandwidth. In this section we will calculate how many events went lost because of this dynamic prescale.
First of all we need to select an unprescaled dataset that contains also a significant amount of PHOTON_DIJET_L2_DPS events as for example the PHOTON_25_ISO trigger. Then looking at the trigger information we can identify a subset $\mathfrak{D}$



(a) $\sum_{L1} E_T$ (b) $\sum_{L1} E_T - P_T^{\gamma}$

Figure 6.14: *The "*emulated*"* $\sum_{L1} E_T$ *(a) and* $\sum_{L1} E_T - P_T^{\gamma}$ *(b) distribution for* $W(Z)\gamma \to \gamma q\bar{q}$ *Monte Carlo events. The trigger thresholds for these quantities are 30 GeV and 20 GeV respectively.*

of these events fulfilling the PHOTON_DIJET_L2_DPS trigger requirements. In a similar way we can identify a subset $\mathfrak{E}$ of events that instead actually *fired* the PHOTON_DIJET_L2_DPS triggers. It is clear that the events that are in $\mathfrak{D}$ but not in $\mathfrak{E}$ have been cut by the dynamic prescale mechanism.
With this method we derived an effective prescale of the PHOTON_DIJET_L2_DPS trigger equal to 1.07 for the period Aug.2003-Sep2003 and a prescale of 1.25 for the period Nov.2003-Jan.2004. The integrated luminosity reported in Table. 6.3 are corrected by these factors.

Chapter 7

Standard Selection

In this chapter we will describe the offline selection cuts we applied on our dataset and we will calculate the total efficiency ϵ of the $W(Z)\gamma \to \gamma q\bar{q}$ signal event selection. As a result, the total number of signal events expected in our sample will be:

$$N_{W/Z\gamma} = \epsilon \cdot \int L \cdot \sigma_{W/Z\gamma}$$

where the $\int L$ is the integrated luminosity of the sample (reported in Table 6.3) and $\sigma_{W/Z\gamma}$ is the signal cross section (reported in sec. 5.5). For explanatory reason we break up the total efficiency in different components (that will be individually addressed in the following sections) and write it as:

$$\epsilon = A_{kin} \cdot \epsilon_{ph} \cdot \epsilon_{jet} \cdot \epsilon_{trg}$$

where A_{kin} is the kinematic and geometric acceptance (sec. 7.1), ϵ_{ph} is the efficiency of the offline photon selection cuts (sec. 7.2), ϵ_{jet} is the efficiency of the jet offline cuts (sec. 7.3) and ϵ_{trg} is the trigger efficiency we have calculated in Ch.6. To assess these efficiencies we will use the PYTHIA Monte Carlo sample we introduced in Ch.5. Corrections for known Monte Carlo deficiencies in reproducing the data will be introduced when needed.

7.1 Event Selection and Signal Acceptance

The offline selection cuts we applied to our dataset can be categorized in 3 groups:

1. Quality Cuts: High quality events are selected requiring at least one primary vertex (sec. 4.5.1) with a $|z_v| < 60$ GeV from the center of the detector, corresponding to the fiducial region of the vertex reconstruction.

2. Photon Cuts: We required the presence of one photon candidate (sec. 4.7.1) with an energy greater of 12 GeV. These selection cuts are summarized in Table 4.1.

3. Jet Cuts: Since we look for the hadronic decays of the W and Z we selected events with at least two reconstructed offline jets [1] with an energy greater than 15 GeV. As jets from QCD are softer (fig. 8.4) than those ones from the boson decays, this energy threshold significantly suppresses the background. Moreover we required the energy of the third jet not to exceed 10 GeV. In this way we reduce the extra jet activity highly present in the QCD background.

To estimate the fraction of signal events that would pass our selection we applied these cuts to the Monte Carlo $W(Z)\gamma \rightarrow \gamma q\bar{q}$ sample. There are two standard ways to calculate the efficiency for a single cut: the efficiency after each successive cut ("*absolute efficiency*"), and the efficiency of the cut assuming the object passes all the previous cuts ("*relative efficiency*"). We adopted the latter definition and in the following for each single cut we will always quote its relative efficiency.
The first cut we applied is the z vertex requirement. Using the Monte Carlo sample we found an efficiency of 0.961. This is in good agreement with the measurement in [41] where minimum bias events were used to derive the longitudinal profile of the $p\bar{p}$ luminous region and an efficiency of 0.951±0.005 was quoted for the $|z|$ <60 cm vertex cut. This agreement was somehow expected because having two energetic jets in the final state of our signal events, our tracking multiplicity is high enough to have a good vertex reconstruction. Therefore the inefficiency of the $|z|$ <60 cut is mainly due to the luminous region acceptance that is essentially a parameter of the beam common to all CDF events.

7.1.1 Photon Acceptance

After the signal Monte Carlo generation we required - using the post-generator filter - the presence of a photon with a P_T >10 GeV and $|\eta|$ <1.2 in the final Monte Carlo particle list (HEPG level). However this does not guarantee that a photon object will be reconstructed at the offline level (it can for instance fall into a crack or fail the E_{HAD}/E_{CEM} criteria). Events with an offline electromagnetic object matching the HEPG photon ($\Delta R(\gamma^{HEPG}, \gamma^{Offline}) < 0.2$) and inside the fiducial region of the CES and CPR detector (see Table 4.1) were selected. Then the $P_T^{\gamma} > 12$ GeV cut was applied. All these criteria define the photon acceptance of our MC sample and the value we found was A_{kin}= 0.377 with 1% of uncertainty.

7.2 Photon Selection Efficiency

To evaluate the relative efficiency of the photon selection criteria we applied *sequentially* the cuts reported in Table 4.1 to the signal events. However other photon studies [42] showed that the Monte Carlo is not able to reproduce correctly the data. This has been attributed to the deficiencies of the detector

[1] In our analysis we used jets reconstructed with a cone radius R=0.7.

simulation (for instance in the tracking and in the hadronic deposition of electromagnetic particles) and in part to the Monte Carlo incapability to account correctly for the underlying events, an important factor in the isolation efficiency calculation. As a consequence, for those cases where the efficiency can be checked with the data we have applied a correction factor to the Monte Carlo efficiency. On the other cases we included the uncertainty of the Monte Carlo simulation in the systematics.
We will now address the efficiency and the associated systematic for each photon cut:

- Had/Cem Ratio: For this cut the Monte Carlo efficiency was 0.945. However the Monte Carlo is known to underestimate this efficiency. Using a random cone in jet events data this efficiency has been measured to be 0.982 with 1% of systematics. The discrepancy between these two values is mainly due to the different activity of the two samples. In our signal events, where we have an hadronic decay of heavy boson, it is indeed more likely for the hadronic energy of the photon towers to be contaminated by some nearby particle shower resulting in a lower efficiency of the E_{HAD}/E_{CEM} cut. For the systematic we used the 1% value used in the data.

- Isolation: The efficiency for the isolation cut was 0.808. This cut is very sensitive to the underlying event that, occasionally depositing some energy (or track) in the isolation cone, reduces the efficiency. Again random cone studies showed the Monte Carlo to be correct within 2% that we set as systematic error.

- Track Cuts: The Monte Carlo efficiency for these cuts was 0.802. In this case, besides the underlying event depositing a track, another inefficiency issue is the photon conversion. When the corresponding track is reconstructed the event is rejected. To correctly predict the photon conversion probability the detector simulation needs to know exactly the amount of material in front of the calorimeter. It was well known that the detector simulation underestimated the material amount of about 4% but this has been fixed in the last version of the offline code (`5.3.3`). Since we used this version to generate our Monte Carlo sample our efficiency accounts for the right photon conversion probability. The tracking reconstruction efficiency in Monte Carlo and data agree within 1% and including the uncertanties on the underlyng event simulation we set a global 2% systematics for this efficiency.

- χ^2 Cut: The efficiency of this cut was 0.996. This is in agreement with other low P_T photon analyses and the random cone study showed that the uncertainty for this efficiency is less than 0.2%.

- 2^{nd} Ces Cluster Cut: The Monte Carlo efficiency for this cut was 0.916. A independent measurement using a $Z \rightarrow ee$ data events [43] found an

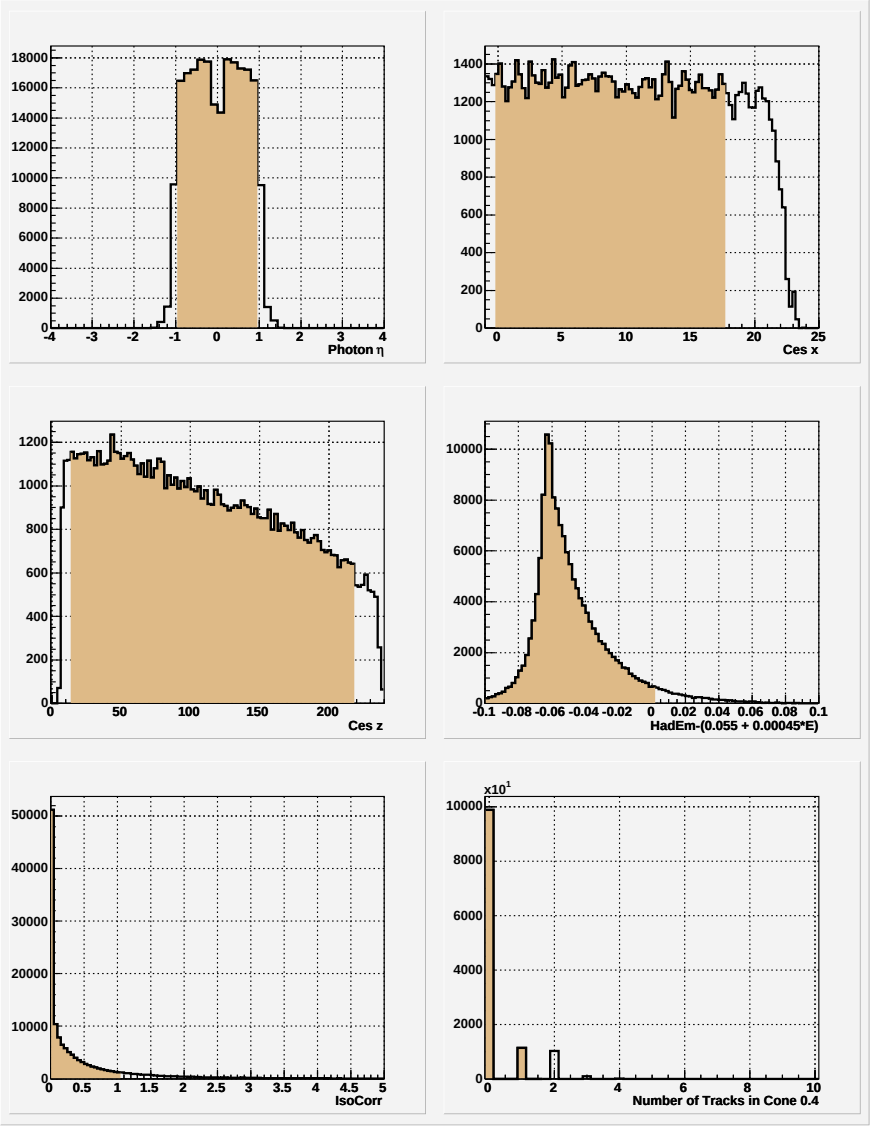


Figure 7.1: *The following photon distributions are shown (from the left to the right, top to bottom): 1) Photon η, 2) CES $|x|$ 3) CES $|z|$ 4) $E_{HAD}/E_{CEM} - (0.055 + 0.00045 \cdot P_T^{\gamma})$, 5) Isolation Energy, 6) Number of tracks pointing to the photon cluster. The region selected by the photon cuts have been highlighted.*

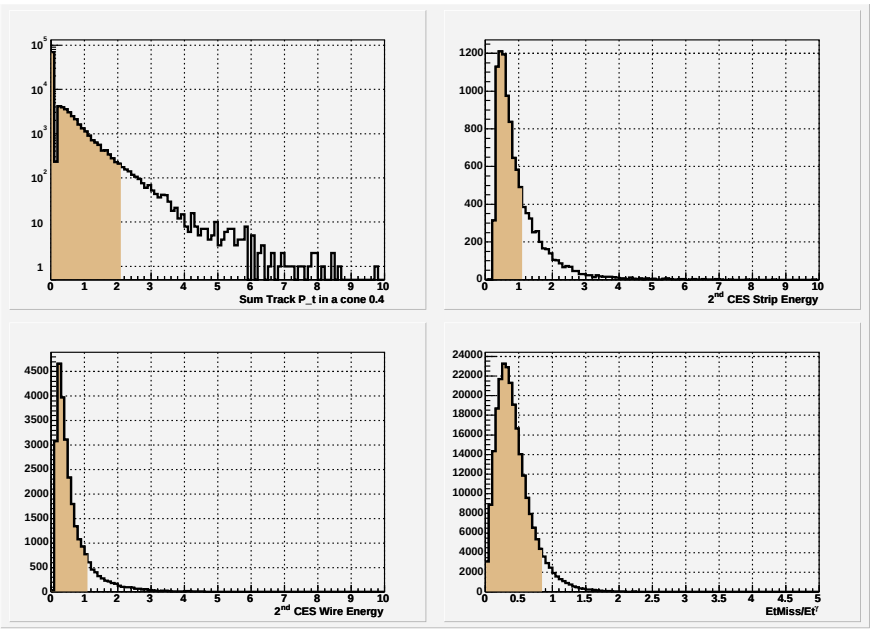


Figure 7.2: *The following photon distributions are shown (from the left to the right, top to bottom): 1) $\sum P_T^{tracks}$ in a 0.4 cone, 2) 2^{nd} CES Strip cluster energy, 3) 3^{nd} CES Wire cluster energy, 4) $\not\!\!E_T/P_T^{\gamma}$. The region selected by the photon cuts has been highlighted.*

efficiency of 0.948. It is known that there are some aspects of the CES reconstruction that are not reproduced in the data (as for instance the presence of bad channels) and therefore a lower efficiency is somehow expected. We decided to trust the efficiency from the data and use the 3% difference as systematic uncertanty.

- Missing E_T Cut: The efficiency found for this cut is 0.90. To set the uncertainty we shifted the 0.8 cut of 0.05 and this resulted in a 2% shifting in the efficiency that we set as systematic.

A summary of the photon cut efficiencies is reported in Table 7.1 along with the corresponding systematic errors.
The total photon identification efficiency resulted to be $\epsilon_{ph} = 0.520$ with 4.6% of systematic error. Fig. 7.1 and 7.2 show the distribution of photon variables before applying the selection cuts.

7.3 Jet Selection Efficiency

After the photon selection we applied to the Monte Carlo sample the jet cuts introduced before. We started selecting the events having a third jet [2] with $E_T^{jet3} < 10$ GeV, then we required a second jet with $E_T^{jet2} > 15$ GeV. As a consequence the first jet will have automatically an $E_T^{jet1} > 15$ GeV. The relative efficiencies of these cuts are 0.546 and 0.828 respectively.
Since we use the corrected energies of the jets, errors in the efficiencies are introduced by the systematic uncertainties of the jet energy corrections. We estimated these errors shifting by 1σ the jet energy scale. It is clear that since a change in the energy scale affects all the jets in the event, we cannot calculate the systematics of each cut separately. They are not independent. Therefore we estimated only the overall systematics for the jet cuts finding a 3.3% systematic error. The resulting total jet selection efficiency was found to be $\epsilon_{jet} = 0.452$ (Table 7.1).
In fig. 7.3 we reported some jet distributions before the selection cuts.

Event Selection Efficiency		
Cuts	*Efficiency* (%)	*Systematics* (%)
z Vertex Cut	96.1	0.5
Acceptance A_{kin}	37.7	1
Had/Cem Ratio Cut	94.5	1
Isolation Cut	80.8	2
Track Cuts	80.2	2
χ^2 Cut	99.6	0.2
2^{nd} Ces Cluster Cut	94.8	3
Missing E_T Cut	90.6	2
Total Photon Efficiency: $\epsilon_{ph} = 0.52$ (4.6% sys)		
Jet 3 Cut	54.6	N/A
Jet 2 Cut	82.8	N/A
Total Jet Selection Efficiency: $\epsilon_{jet} = 0.452$ (7.4% sys)		

Table 7.1: *Summary of the event selection efficiency.*

[2] The jets are sorted according to their corrected energy

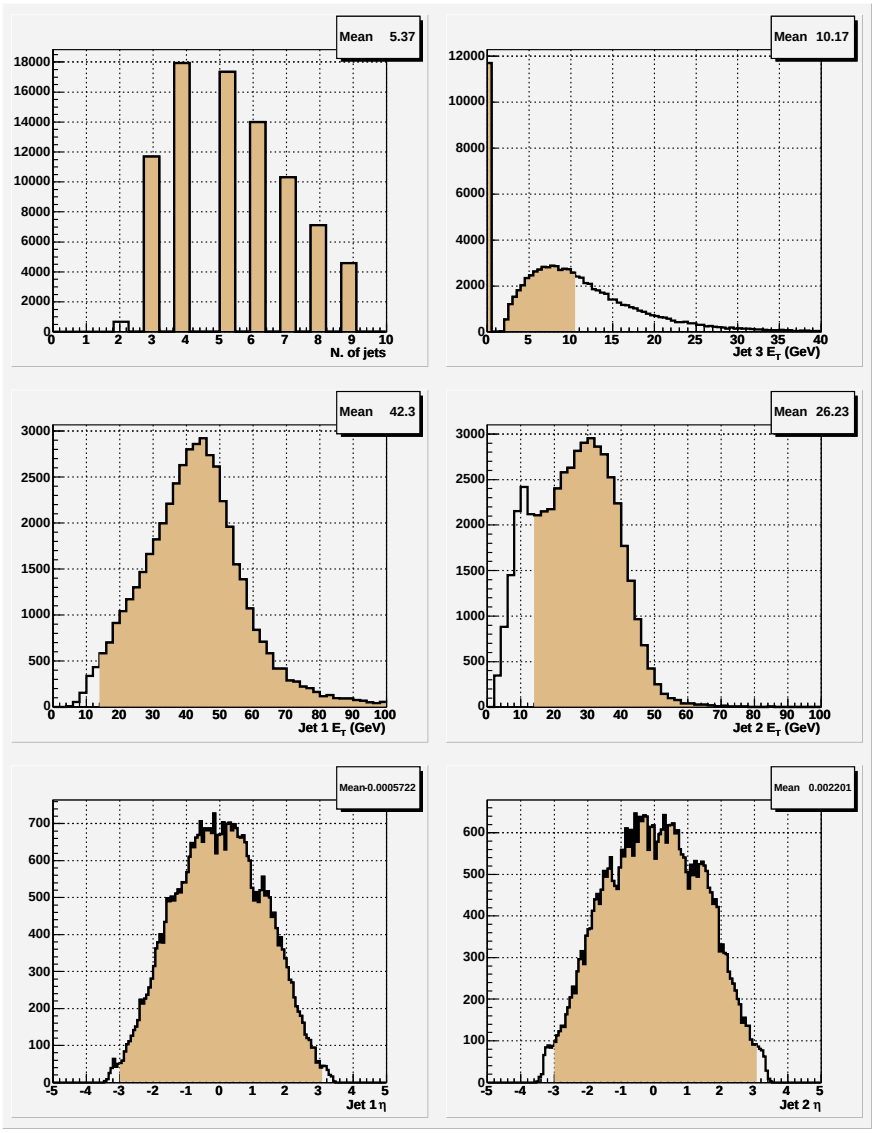


Figure 7.3: *Distributions of jet quantities before the selection (from the left to the right, top to bottom): 1) Number of reconstructed jets (including the jet matching the photon); 2) E_T of the third jet; 3) E_T of the first jet; 4) E_T of the second jet; 5) η of the first jet; 6) η of the second jet. The region selected by the jet cuts has been highlighted.*

7.4 Summary

Now we are ready to combine all the selection efficiencies together and add as well the trigger efficiency calculated in the previous chapter.
For the event selection we have:

$$\epsilon_{sel} = \epsilon_{zvtx} \cdot A_{kin} \cdot \epsilon_{ph} \cdot \epsilon_{jet} = 0.081$$

with 5.7% systematics. For the trigger efficiency there are two different values we need to consider:

$$\begin{cases} \epsilon_{trg}^{DPS} = \epsilon_{ph}^{trg} \cdot \epsilon_{jet}^{trg} = \epsilon_{dj10 \to 10iso} \cdot \epsilon_{10iso} \cdot \epsilon_{sumet}^{DPS} \cdot \epsilon_{seed} = 0.838 \pm 0.008 & \text{with 1.2\% sys.} \\ \epsilon_{trg}^{dj12} = \epsilon_{ph}^{trg} \cdot \epsilon_{jet}^{trg} = \epsilon_{dj12 \to 10iso} \cdot \epsilon_{10iso} \cdot \epsilon_{sumet}^{dj12} \cdot \epsilon_{seed} = 0.758 \pm 0.01 & \text{with 1.5\% sys.} \end{cases}$$

where ϵ_{trg}^{DPS} and ϵ_{trg}^{dj12} are for the `PHOTON_DIJET_L2_DPS` and `PHOTON_DIJET` trigger respectively.
Multiplying the two contributions we can finally find the total efficiency of our sample:

$$\begin{cases} \epsilon_{DPS} = \epsilon_{trg}^{DPS} \cdot \epsilon_{sel} = 0.071 & \text{for the } \texttt{PHOTON\_DIJET\_L2\_DPS} \\ \epsilon_{dj12} = \epsilon_{trg}^{dj12} \cdot \epsilon_{sel} = 0.065 & \text{for the } \texttt{PHOTON\_DIJET} \end{cases}$$

with a 8.9% systematic uncertanty in both cases.

7.5 The Signal Dijet Mass Distribution

With the final subsample of $W(Z)\gamma \to \gamma q\bar{q}$ events passing all the selection criteria we plotted the dijet mass distribution of the two leading jets in order to reconstruct the W/Z mass peak as it should appear in the data. Using the Monte Carlo events we have also the luxury of knowing if the boson is a W or a Z and in this way we are able to report in the same plot the two components of the peak. Some non-gaussian tails in these individual boson distributions are clearly visible. To understand the reasons of these tails we have to bear in mind that even though these are $W(Z)\gamma \to \gamma q\bar{q}$ events, the two quarks from the boson decays are not the only partons in the final state. Partons (mainly gluons or quarks from gluon splitting) from the initial state radiation (ISR) can generate jets stiff enough to be promoted to one of the first two positions in the energy ranking. In this case the dijet invariant mass - made with a spurious ISR jet - has nothing to do with the W/Z mass [3]. Another possibility is instead that one of the two leading jet is a gluon [4] from the final state radiation (FSR), *id est* radiated by one of the quarks from the boson decay. In this case only part of the boson energy is carried by the first two jets and as a consequence the dijet invariant mass is lower than the expected value. It is clear that in order to reconstruct at the best the W/Z peak it is essential to reduce the occurrence

[3] These events are typically responsible for the high tail of the dijet mass spectrum.
[4] Or more rarely a quark from the splitting of such gluons.

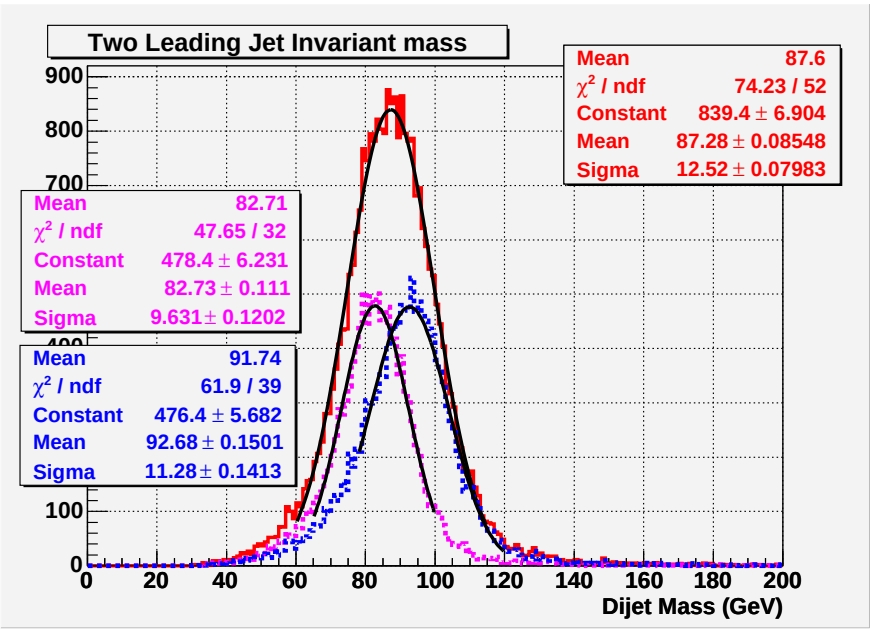


Figure 7.4: *The dijet invariant mass (red full line histogram) of the two leading jets in the final selected $W(Z)\gamma \to \gamma q\bar{q}$ Monte Carlo events. The individual contributions from the W (dashed purple) and Z (dashed blue) bosons are shown. All the distributions are fitted with a gaussian and the result of the fits reported in the insets.*

of these cases at the minimum. The 10 GeV cut on the third jet was meant exactly for this purpose. In the former case, since the second (or the first) jet is from ISR, one jet from the boson decay needs to be the third jet (or a higher jet). But usually the jets from the W(Z) are rather energetic allowing the 10 GeV to strongly suppress this sort of events. In a similar fashion this cut also suppresses the events with a stiff jet from the FSR. Nevertheless soft FSR jets (with $E_T < 10$ GeV) can still be present in our selected event. The consequences of this soft gluon radiation are recognizable in the low mass tail noticeable in the W and Z peak in fig. 7.4 [5]. Lowering further the third jet threshold would definitely reduce this effect but at the price of a worse efficiency (that with our cut is already 45%). All in all we gauged the 10 GeV threshold to be the best compromise between FSR effect and efficiency [6].
Anther approach to recover in part the FSR radiation is the use of a large jet

[5] The reason for the longer tail in the Z peak with respect to the W peak can be attributed to the 15 GeV cut on the second jet energy. Having a higher mass, the jets from the Z are more energetic and therefore more likely they can radiate a gluon and *still* remain enough energetic to pass the 15 GeV cut.

[6] Also the suppression of the QCD background needs to be accounted, favoring a low third jet threshold.

cone size. We have shown in another study [44] that for the identification of a W boson decaying into a pair of jets the cone radius R=1.0 is optimal to recover the soft FSR emitted not far from the center of the jet. However, when also a photon is present in the event, wide cones are reconstructed with a lower efficiency because more likely they overlap with the photon electromagnetic cluster. Our choice of a cone 0.7 is again a good compromise.
Another important feature of the dijet mass distribution is its shape. As it can be noticed looking at the plot, the distribution is well fitted by a gaussian if we exclude the small tails outside the 60-120 GeV region. This is true in spite of the fact that the W and Z components are not gaussian caused by the larger tails present in particular at low mass. A quantitative way to arrive at the same conclusion is to make the fit in a smaller region around the peak to exclude the possible tails and see if there is a change in the fitting outcome. We checked that, fitting the distribution in the 70-110 GeV and 80-100 GeV region, and we found basically the same results (inside the statistical errors). For this reason, we defined the shape of our signal as a single gaussian - with $\mu = 87.28$ and $\sigma = 12.52$ - instead of using the sum of the W and Z distributions whose shapes we do not know how to parametrize anyway.
The final comment on the plot in fig. 7.4 is about the dijet mass resolution. Using the fit results of the W and Z peaks we found a $\delta(M)/M$ resolution around 12% for both peaks. This is consistent to the known number of the CDF dijetmass resolution when only the standard jet corrections are applied.

Chapter 8

Background

8.1 Introduction

The selection cuts we discussed in the previous chapter provide a sample of events with a well identified candidate photon along with (at least) two reconstructed jets. Although the cuts have been tailored on $W(Z)\gamma \to \gamma q\bar{q}$ events, there are several physics processes that can produce such a final state. We conveniently grouped them in two main categories: (1) the γ + two jets production and (2) the three jets production where one of the jets is identified as a candidate photon ("*jet fake*"). Of course in the former group is included the resonance production of the W/Z decaying in two jets, the topic of this dissertation. These two contributions are roughly equally represented in our final sample with the ratio depending on the photon energy. A measurement of the prompt photon fraction in photon candidate events was provided in fig. 4.10 and, since the jet selection does not alter the relative composition of the sample, the same plots can be interpreted as the ratio between the two aforementioned set of processes. It can be noticed that for a P_T^γ between 12 and 18 GeV (where most of our sample lies) about 60% of the photon candidates are acutally jet fakes. However - as we mentioned several times already - these events can *completely* be removed using the photon background subtraction technique described in sec. 4.8.1. This highlights how essential this tool is for this analysis. In fact it is able to remove more than 60% of the background (the three jet contribution) without reducing the signal efficiency, a big enhancement in the purity of the sample.
After that the only background left is the γ + 2 jets production and in the following sections we will discuss in details its properties and production mechanism .

8.2 γ + 2 jets Production

From a theoretical point of view, the presence in the final state of one photon and two partons (quark or gluon) is addressed with a $O(\alpha\alpha_s^2)$ NLO calculation of

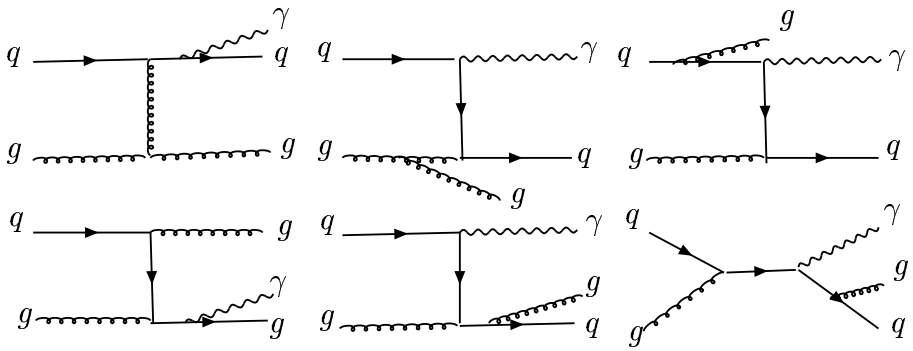


Figure 8.1: *Feynman diagrams for some γ + 2 jets production mechanism. The first two diagrams on the left are examples from the bremsstrahlung contribution.*

the inclusive prompt photon production [46]. Some Feynman diagram examples of these processes are shown in fig. 8.1. Although the simple photon coupling makes relatively easy the NLO calculation for the direct production, the cases with the photons emitted collinear with a quark (usually called "*bremsstrahlung contribution*" or "*collinear fragmentation*") add some divergences that need to be carefully managed. Indeed, although the rate for the photon radiation is small compared to the gluon radiation, the dijet-production cross section is sufficiently large that the bremsstrahlung contribution is predicted to be of the 1same order of the direct one.

This is only one of the theoretical issues related with the NLO inclusive photon production calculation and lengthly debated in the literature. However since we are interested in these events only as background of our analysis we limited our study to their kinematics properties. In particular we want to compare them with those of our signal events in order to have some insights on how to increase the purity of our sample. To perform this study we decided to generate a Monte Carlo sample of γ + 2 jets events.

The first thing we had to decide was how to produce this Monte Carlo sample. The question is not trivial because the jets in the final state can be produced not only in the hard scattering (as in the cases depicted in fig. 8.1) but also during the parton shower. However - as discussed in sec. 5.1 - for computation purposes the hard scattering and the fragmentation are completely separated processes that often are even handled by different Monte Carlo programs. For this reason a γ + 2 jet production made with a matrix element Monte Carlo will include the cases of both jets produced in hard scattering but not γ + 1 jet events with other jets produced at the fragmentation level. [1]. The only neat way to have a trustful γ + 2 jet sample would be to generate an inclusive photon sample with a NLO Monte Carlo and apply the jets requirements *after*

[1] Here we totally ignore the further complication of events with the *photon* generated during the fragmentation. The tight isolation cut justifies this decision.

the full simulation [2].
Unfortunately this approach is not practically feasible because the fraction of photon events with two jets in the final state is very small and it would require the generation of a huge amount of inclusive photon events before extracting a significant sample of $\gamma + 2$ jets.
In spite of the aforementioned problems we decided anyway to use a matrix element Monte Carlo program for the $\gamma + 2$ jets event generation.. Considering that we are interested only in a limited number of properties of these events, that we are selecting two energetic jets and that we do not seek for an high level of accuracy for the cross section, we found this approach the most practical postponing any judgment on the usefulness of the sample until the scrutiny of the final state object distributions. We therefore generated 2 million $\gamma + 2$ jets events using MadGraph as a matrix element calculator followed by the PYTHIA parton shower and the CDF detector simulation, exactly with the same procedure used in the $W(Z)\gamma \to \gamma q\bar{q}$ signal production described in sec. 5.2.2. To be consistent we used the same parameter configuration (PDF, q^2, etc...) and the same cuts: $P_T^\gamma > 6$ GeV, $|\eta^\gamma|$ <3 GeV and $\Delta R(\gamma, q) > 0.4$ at the generation level, P_T^γ >10 GeV and $|\eta^\gamma| > 1.2$ GeV at particle level. The resulting LO cross section was about 284 nb. The comparison with the corresponding LO signal cross section (15.42 pb) is stunning, giving an idea of the huge amount of background one has to deal with in order to dig out a possible hadronic W/Z peak [3].
Another way to visualize the amount of the $\gamma + 2$ jets background contribution is reported in fig. 8.2 where the W and Z resonances are barely noticeable from the falling QCD spectrum.
The first thing we looked at with the new sample are the final state observable distributions at parton level. In particular we plotted the P_T, η and the distance ΔR of the three final partons (fig. 8.3). It is interesting to notice the enhancement in the $\Delta R(\gamma, q)$ distribution for small value (close to the 0.4 cut) due to the bremsstrahlung contribution. A similar enhancement is clearly visible also in the $\Delta R(q_1, q_2)$ distribution when, instead of a photon, a gluon is emitted.
These distributions can be directly compared with those for the $Z\gamma \to q\bar{q}$ events (fig. 5.3) and for the $W\gamma \to q\bar{q}$ (fig. 5.4). Such a comparison suggests that the differences in the kinematics properties of the signal and the background can be exploited to discriminate the two samples, if it is done in a proper and systematic way. This is exactly the approach we pursued in this analysis as we will show in details in the following chapter.

[2] This problem is somehow similar to the $W + n$ jets production where a greater (or less) number of jets can be found in the final state when exactly n are produced in the hard scattering.

[3] Actually for a fair comparison we should calculate the $\gamma + 2$ jets cross section only in the phase space where the W/Z lie.

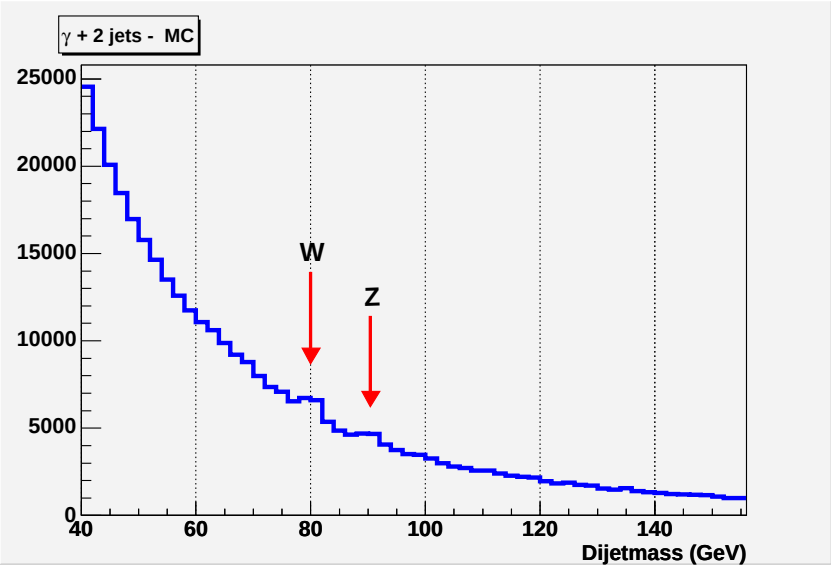


Figure 8.2: *Mass distribution of the two final partons in γ + 2 jets Monte Carlo events. It includes the W and Z resonance production as highlighted in the plot.*

8.3 γ + 2 jets Data-Montecarlo Comparison

To complete the Monte Carlo validation we need to check if the $\gamma + 2$ jets sample is able to reproduce correctly the data distributions. This is of paramount importance for the success of the signal/background discrimination technique we employed in this analysis. Indeed, the advanced selection we are going to apply in the next chapter is completely based on the comparison of the distributions generated with a signal and a background template sample in order to exploit their differences in *shape* to gain discrimination power. It is mandatory therefore that the MadGraph $\gamma + 2$ jets Monte Carlo sample reproduces the shape of the data distribution correctly, if we want to use it as a template for the background [4]. From fig. 8.4 to fig. 8.6 a direct comparison between the data collected with the inclusive `PHOTON_15_TIGHT` trigger and the $\gamma + 2$ jets Monte Carlo sample is reported. In order to have a clear picture of the possible differences between data and Monte Carlo we compared several distributions of final state variables. As it can be seen, in general the agreement is good for basically all the variables we plotted. However the Monte Carlo is not able to reproduce with the necessary precision the *shape* of these distributions. We did not investigate further the

[4] We need to remember that the W and Z boson resonances (our signal) are embedded in the $\gamma + 2$ jets we generated. This means (1) that it contains exactly the same processes that are in the data and (2) that the signal contribution needs to be removed if we want to use this sample as a *background only* template. It can be easily done (with good approximation) changing the setting of the MadGraph input parameters.

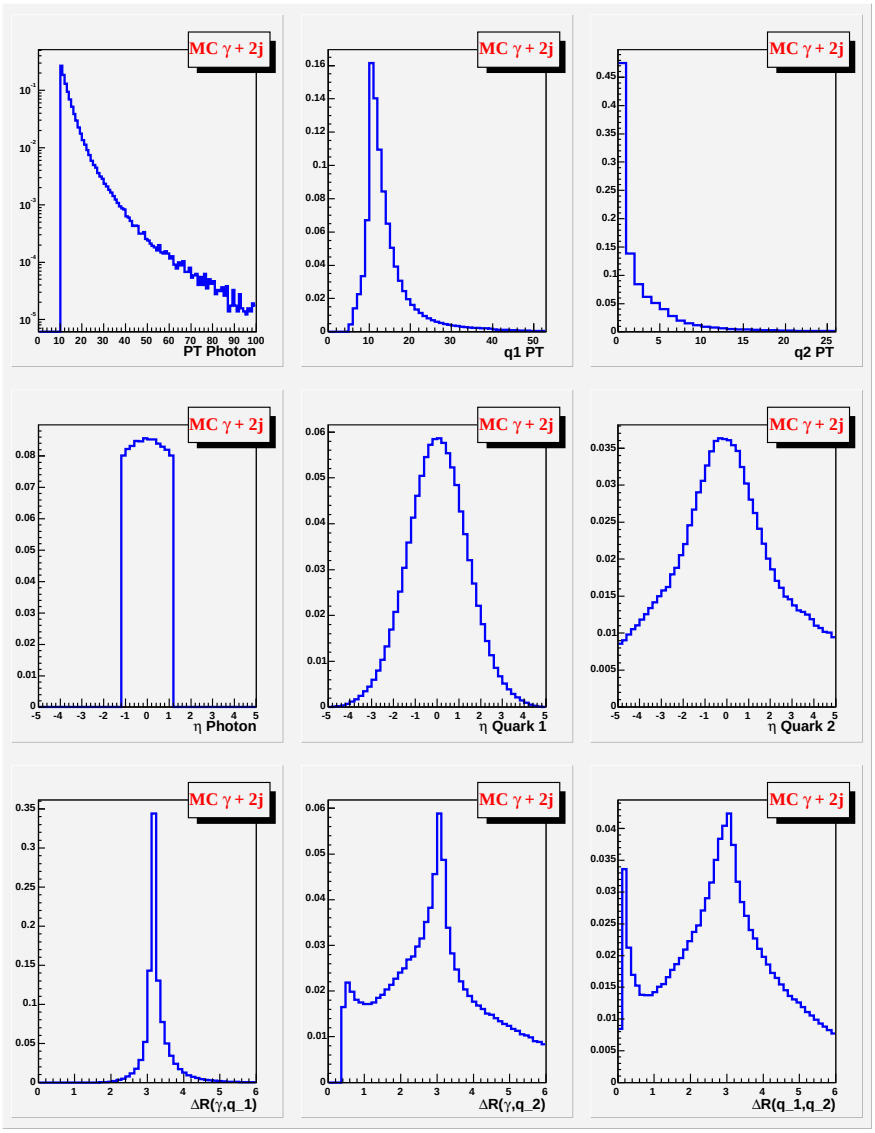


Figure 8.3: *Final parton kinematics distributions of* MadGraph *Monte Carlo* $\gamma + 2$ *jets events. The momentum, pseudorapidity and the distance between each other* (ΔR) *are shown. On the* $\Delta R(\gamma, q_2)$ *distribution is visible the effect of the bremsstrahlung contribution. The effect of the gluon radiation emitted by one of the quark is visible in the low value of the* $\Delta R(q_1, q_2)$ *distribution.*

reason of these discrepancies, but because the differences are small the use of a LO instead of NLO Monte Carlo could have had some crucial influence - besides the intrinsic generation problems discussed at the beginning.

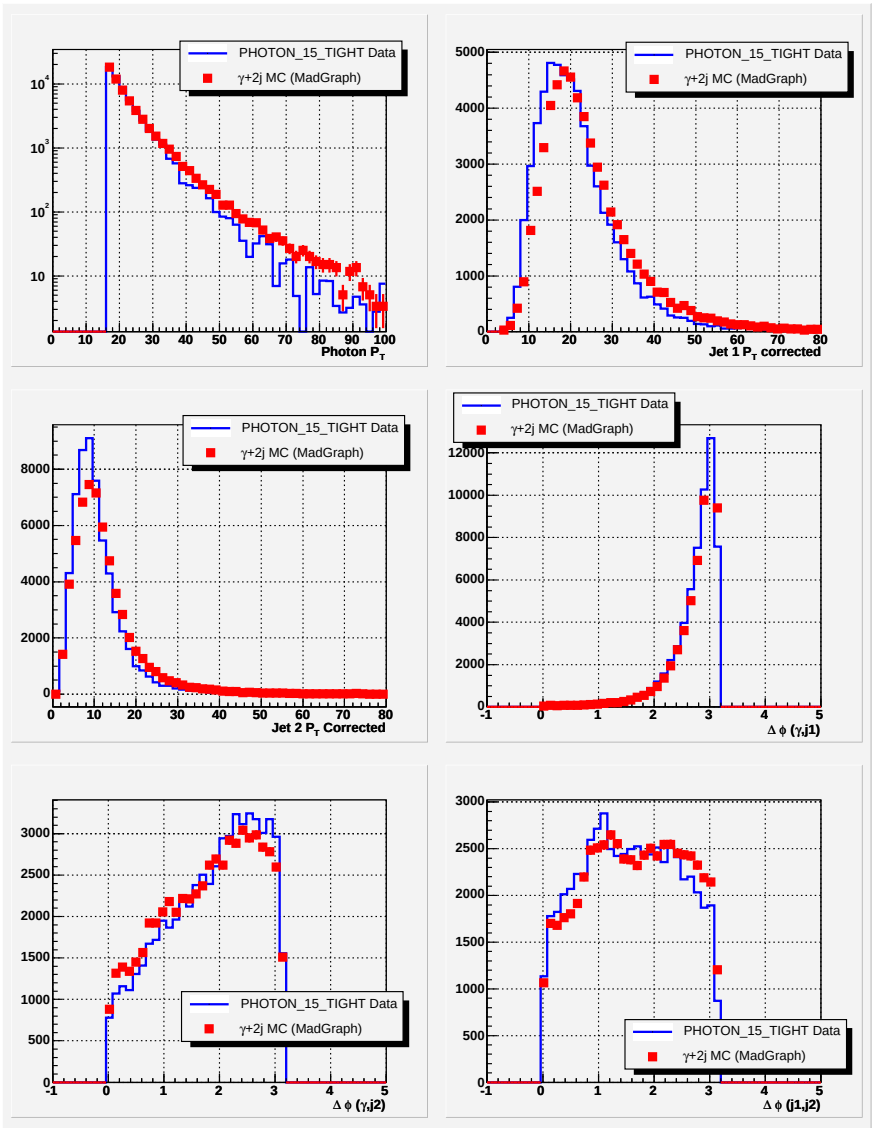


Figure 8.4: *Comparison of data (*PHOTON_15_TIGHT *trigger) and Monte Carlo (*MadGraph*) γ+2jets events. Only the photon selection cuts with a P_T^{γ} threshold of 16 GeV are applied. For the data the photon background subtraction are applied as well.*

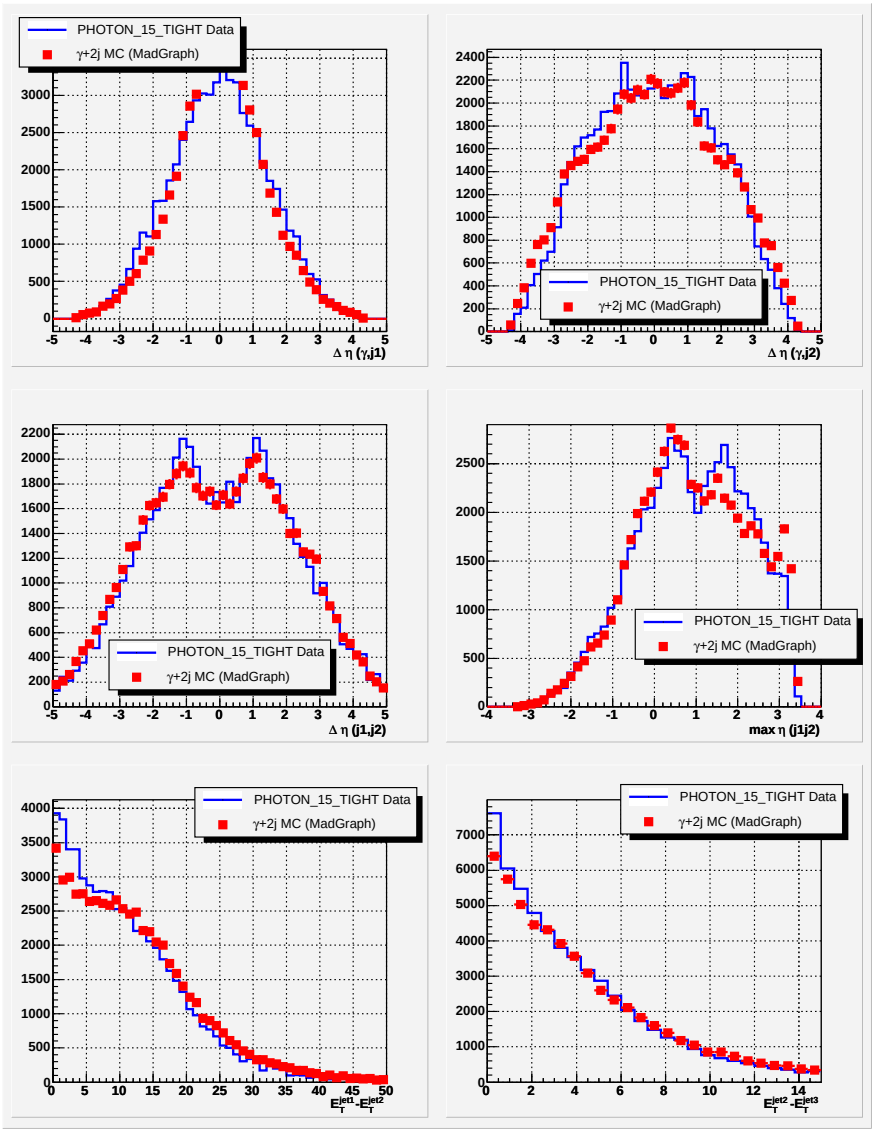


Figure 8.5: *Comparison of data (*PHOTON_15_TIGHT *trigger) and Monte Carlo (*MadGraph*) γ+2jets events. Only the photon selection cuts with a P_T^{γ} threshold of 16 GeV are applied. For the data the photon background subtraction are applied as well.*

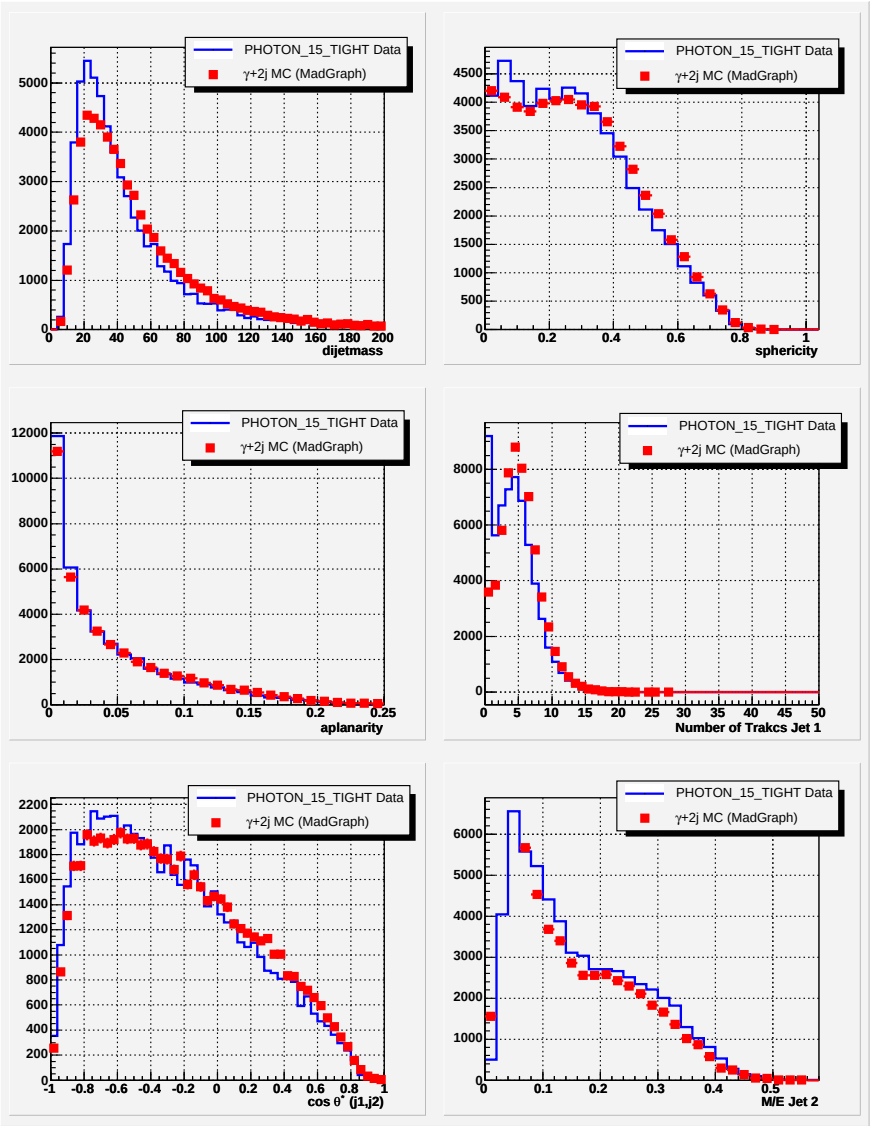


Figure 8.6: *Comparison of data (*`PHOTON_15_TIGHT` *trigger) and Monte Carlo (*MadGraph*) γ+2jets events. Only the photon selection cuts with a P_T^{γ} threshold of 16 GeV are applied. For the data the photon background subtraction are applied as well.*

8.4 Conclusions

After the photon background subtraction the only background left in our sample is the QCD $\gamma + 2$ jets production. Several possible hard processes may produce two jets in the final state associated with a photon and among them there are, of course, the W and Z production. Unfortunately the *MadGraph* $\gamma + 2$ jets Monte Carlo sample we generated does not reproduce the data with the accuracy needed for our purpose and it therefore cannot be used as a template for the background. However since the number of $W(Z)\gamma \to \gamma q\bar{q}$ signal event are really small in this sample (as it is shown in fig. 8.2) it is reasonable to assume that their presence does not have a significant impact on the $\gamma + 2$ jets distributions. Under these circumstances we judged that it is perfectly legitimate to use directly the *data* as a background template. On the other hand it is much better to have a template with a negligible amount of spurious events than having a template with only the "right" events but that does not reproduce correctly their distributions.
In the following chapter we will finally introduce the advanced selection we adopted in this analysis to try to identify the $W(Z)\gamma \to \gamma q\bar{q}$ signal events enhancing the purity of our sample.

Chapter 9

Advanced Selection

9.1 Expected Event Yield

Given the selection criteria we described in Ch.7 and the corresponding efficiencies, we can now calculate how many signal events are expected in our data sample, and compare them with the number of events we actually have in the data inside the dijet mass region of the W/Z resonance peak. The dijet mass spectrum for the signal and for the data sample are shown in fig. 9.1 where the mass window used for the the counting has been conveniently highlighted. The signal peak has been already discussed in sec. 7.5. The data distribution was obtained with the same selection criteria of the signal and applying the photon background subtraction.

Using these distributions, the total integrated luminosity of the data sample, the selection efficiency and the signal cross section we found $S = 227$ expected signal events and $S + B = 75675$ events in the data [1]. Quoting the event yield per picobarn we have about 1.2 ev/pb for the signal and about 411 ev/pb for the data with a signal-over-background ratio of 1/333. To assess the sensitivity of a peak search analysis is however more useful to calculate the significance, defined as $S/\sqrt{S+B}$. In fact, a signal peak can be identified only if it emerges well above the statistical fluctuations (proportional to $\sqrt{N}$) of the background. With a small sample even a good S/B does not guarantee a signal peak to be distinguishable from the large statistical fluctuations. In a big sample instead, the tiny statistical error allows the identification even of a small excess, in spite of a poor S/B.

In our case $\frac{S}{\sqrt{S+B}}$ is 0.8. Since this number depends on the amount of collected data, it is interesting to see how it improves as a function of the integrated luminosity (fig. 9.2). As it can be seen, even with 1 fb^{-1} of data the significance of our analysis is too poor for a successful identification of the W/Z peak [2]. Since

[1] We do not quote here any errors because we reported these numbers just as a reference to be compared with those ones provided in the following sections.

[2] It is not possible to establish a value of the significance for which a peak can be seen. As a rule of thumb at least a significance around 2.5-3 is needed for a a peak to start emerging.

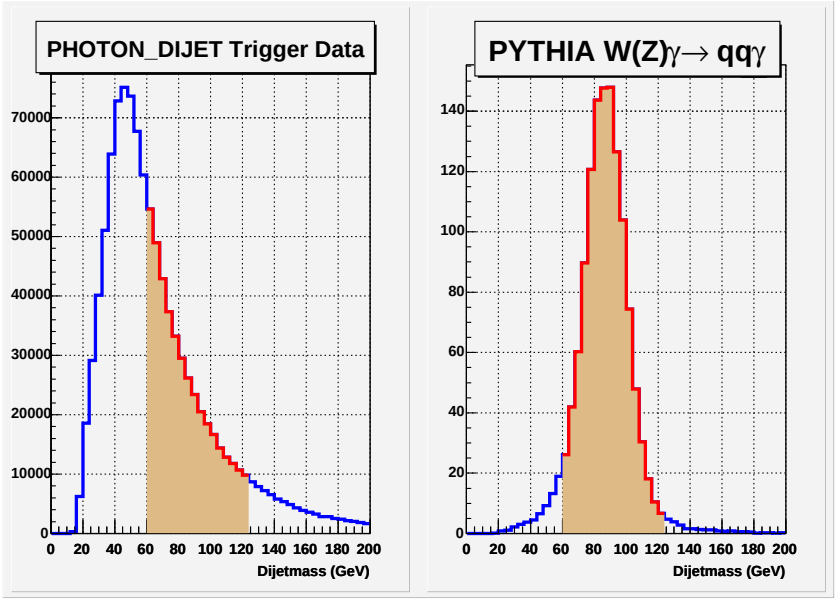


Figure 9.1: *The dijet mass distributions of the two leading jets for the data (left) and Monte Carlo (right) after applying all the selection criteria. For the data the photon background subtraction was applied as well. The region we used to count the number of events is highlighted.*

our selection gives such a poor significance, we tried to enhance our signal sensitivity applying an advanced event selection (on top of the "standard" selection we already employed) performed with a neural network. The description of this new selection and its sensitivity in the W/Z peak search will be discussed next.

9.2 Event Selection with Neural Network

The Artificial Neural Network (ANN) method offers a simple way of combining the information from many kinematic and geometry variables, increasing in principle our potential to resolve the signal in the data. The simple requirements we implemented so far - two jets above threshold and a small extra jet activity - just help to get rid of some manifest background, but overall they have a very mild discrimination power. Not surprisingly - given the huge background we have to deal with (see Ch.8) - the significance after these cuts turned out to be very small. However, as mentioned at the end of sec. 8.2, the event observable distributions of the signal and of the background are different in many ways. To take full advantage of that we have studied a wide variety of variables trying to identify a set of selection criteria capable of a strong signal/background separation. We concluded that none of them were able to perform an event-by-event

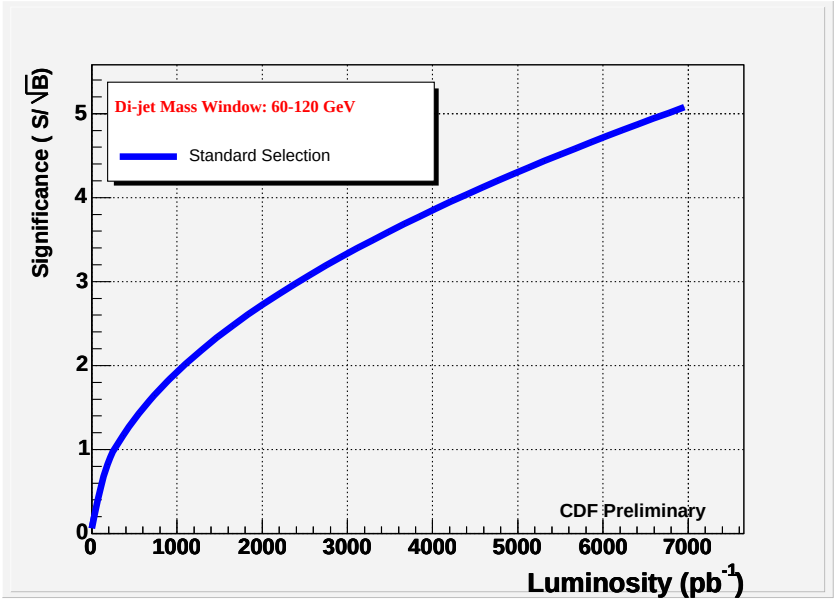


Figure 9.2: *The significance of our selection as a function of luminosity along with the S/B ratio. The value at L=184 pb is the significance of our dataset.*

discrimination although the impression of subtle differences remains.
This is a classic example where *multivariate techniques* were proved very powerful. In fact they are able to exploit information from several sources *simultaneously*, and capable of identifying correlations in multiple dimensions. Moreover, using an arbitrary number of hyperplanes they have the ability to select regions of the variable phase space with complex boundaries of virtually any shape.
The ANN is one kind of these multivariate techniques. In reality it is a very general tool that serves multiple purposes, depending upon its architecture. Because of the extent of the subject we cannot present here a treatise on neural networks [3]. What we can say here is that the main characteristics of the NN are the non-linearity (in contrast with the Fisher Discriminant [48] method) and the use of learning (or training) processes. In the following sections we will describe the architecture of the network we have used for our analysis along with a brief introduction of the NN terminology.

9.3 Neural Network Model

The basic component of any NN is the *neuron* (or node, or unit). It receives input either from the original data provided by the user or from the output of

[3] The interested reader can benefit the widely available numerous resources on the subject (see for instance [47]).

other neurons in the network. Each input comes via a connection that has a strength (or *(synaptic) weight* [4]) and to each neuron is associated a threshold value. The weighting sum of the inputs with this threshold subtracted form the *activation* signal of the neuron (also known as *post-synaptic potential*). Such a signal is then passed to an *activation function* - that is a step function [5] - that eventually gives the state (or *output*) of the neuron; *fired*(1), or *not fired* (0). This describes an individual neuron. How the neurons are connected together in a network is established by its architecture.
In a network there are input neurons (carrying the value of the variables of interest), output neurons (forming the prediction) and hidden neurons [6] (playing an internal role). To connect the neurons we employed a *feed-forward network*: signal flow from inputs, forward through any hidden neurons, eventually reaching the output units [7]. This is a standard choice when NN are applied to signal/background separation problems like ours. In a simple network the neurons are arranged in a distinct layered topology (*perceptrons*) where unidimensional connections runs only between node in consecutive layers (fig. 9.3(a)). In our case we used a *fully-connected* multi-layered perceptrons (MLPs) with each neurons connected to *all* the neurons in the preceding layer through weights and thresholds. They are the parameters of the networks that can be chosen to model functions of almost arbitrary complexity, varying the number of layers and the number of units in each layer. Since one hidden layer is usually enough to approximate any continuous function [8] , we adopted a 3 layer perceptron network with one input, one hidden and one output layer.
The number of input nodes (and which variables to use) will be discussed in details later along with the number of hidden nodes. For the output, since we are dealing with a *classification* problem (determination to which of a number of discrete classes a given input belongs), the output should be a single nominal value. Having only one homogeneous background ($\gamma + 2$ jets events) one output node is enough for us. During the training we will set its answer to 1 for signal and 0 for background events.
When the network is used, after placing the variable values in the input units, the hidden and output layer units are progressively executed. Each of them calculates its activation value (taking the weighted sum of the unit outputs in the preceding layer and subtracting the threshold), passes it through the activation function and produce the output. In our network for example, calling g the activation function, the output H^h of the h^{th} hidden node is (fig. 9.3):

$$H^h = g(\sum_{inputs} w^{hn} x_n - \mu^h)$$

[4]The weights can be negative, which implies that the synapse has an inhibitory rather than excitatory effect on the neuron.

[5]In reality this function need to be differentiable and "S-shaped" functions are typically used. In our network we used the function tanh.

[6]They are *hidden* because they have no connection to the outside world.

[7]This means that in this structure there is no connection from later to earlier neurons as in *recurrent networks* used for example in track-finding algorithms.

[8]This follows from the "universal approximation theorem", see...

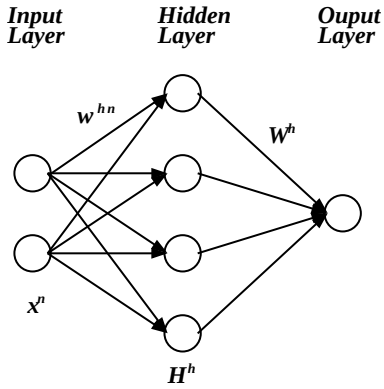


Figure 9.3: *Schematic view of a neural network with one hidden layer and one output node. Each neuron is connected with all the neurons in the previous and following layer. The synaptic weights of the hidden nodes are w^{hn}, where h is the index of the hidden node and n is the index of the input node. The weights for the output are W^h.*

where x_n is the n^{th} input, w^{hn} the corresponding weight for the h^{th} hidden unit and μ^h is the threshold. In similar way the one value output is:

$$O = g(\sum_{hidden} W^h H_h - \lambda)$$

where now W^h are the output weights and λ the output threshold. Thus, once the parameters of the network ($\mathbf{w}$, $\mathbf{W}$, μ and λ) have been set, the output depends only upon the input values $\mathbf{x}$. The determination of these parameters - done during the construction of the network - is the role of the *training algorithms*, the feature that differentiate at the most the ANN to other multivariate techniques.

Training Multilayer Perceptron

Once the number of layers - and the number of unit in each layer - has been decided, the network weights and thresholds must be set such as to minimize the prediction error. This is done running the events of a training sample through a particular configuration of the network and comparing the actual generated outputs with the desired (or target) outputs. The differences are combined together by an error function (for instance the sum squared error) to give the network error. Therefore for each value of the network parameters an error can be determined. If N is the total number of parameters and P_i are the parameter values, all these errors $e(p_1, \ldots, p_N)$ can be visualized as a *surface* in a N+1 dimensional space. The goal of the network training is then to find the minimum of this multi dimensional surface that represents the smallest error the network can make. Unfortunately the error surface can be very complex, characterized by a number of nasty features (like local minima, plateaus, saddle-points,etc..)

and it is not possible to analytically determine the global minimum. What the training algorithms do is essentially an exploration of the error surface - starting from an initial random point - and moving downhill along the gradient of the surface and eventually stop in a low point, which hopefully is a global minimum. Several algorithms have been conceived to perform such a task. In particular we used the *back-propagation* algorithm [9] whose details cannot be described here. Roughly speaking this algorithm uses a "sequence of moves" along the surface gradient with a step size proportional to the slope and to a constant called *learning rate* (that is provided by the user). Also a *momentum term* is included to encourage movement in a fixed direction and move rapidly over a flat spot. However the most important parameter for us is the number of *epochs*, *id est* the number of steps. During the training of our sample we were careful to set this number high enough to allow the algorithm to actually reach the global minimum.
In this analysis we employed the JETNET [49] software package to construct our NN along with the available ROOT interface that allow an easy access to the tunable parameters the user can optimize. In the following sections we will describe the studies made to build a NN suitable to our analysis and we will report the performance of such a network in our data.

9.3.1 Training Samples

As we explained in the previous section, the NN weights and thresholds are determined through a learning procedure where the network is fed with template samples whose (desired) output is known. In our case we need two templates, one for the signal and the other for the background. For the signal we have used of course our PYTHIA $W(Z)\gamma \to \gamma q\bar{q}$ events For the background, since it was not possible to generate a reliable Monte Carlo sample, we have used the data. It is clear that with a $1/343$ signal over background ratio, the data distributions give an excellent approximation of the background behavior.

9.3.2 The NN Inputs

The choice of the NN input variables is by far the most important part in the construction of the network and therefore a considerable effort was put into the determination of such variables. First of all we performed a pre-selection of variables we gauged to have a good separation power. Since there is no rule to make such a selection we were guided by our intuition only and by theoretical considerations of the processes involved. In this phase we were very careful to not peak variables able to sculpt the dijet mass distribution. We need to remember that whatever selection we do, at the end the low mass sideband of the dijet mass spectrum must not disappear. After the scrutiny of several observable combinations we selected 19 variables as possible NN inputs:

[9] Other algorithms are for example the conjugate gradient descent, the quick propagation, the Delta-Bar-Delta, etc..

1. $\Delta\eta(J_1, J_2)$;
2. Number of tracks inside a 0.5 cone [10] around the two leading jets;
3. M/E of the second jet ($M = \sqrt{E^2 - P^2}$);
4. Maximum η of the two leading jets;
5. "Intrajet Energy Ω" defined as $\Omega = \left(\sum_T E_T - E_T^{jet1} - E_T^{jet2} - E_T^{\gamma}\right)/\Delta L$ where $\sum_T E_T$ is the E_T scalar sum of the calorimeter towers in the region $(\eta^{DW} - 0.3) < \eta < (\eta^{UP} + 0.3)$ and $\Delta L = \eta^{UP} - \eta^{DW} + 0.6$ with $\eta^{DW} = \min(\eta^{jet1}, \eta^{jet2}, \eta^{\gamma})$ and $\eta^{UP} = \max(\eta^{jet1}, \eta^{jet2}, \eta^{\gamma})$. The energies of the photon and the two jets are *uncorrected*;
6. $\left(E_T^{jet1} - E_T^{\gamma}\right)/\left(E_T^{jet1} + E_T^{jet2} + E_T^{\gamma}\right)$;
7. $\Delta\Phi(j_1, j_2)$;
8. $\max\left(\Delta\Phi(j_1,\gamma), \Delta\Phi(j_2,\gamma)\right)$;
9. $\min\left(\Delta\Phi(j_1,\gamma), \Delta\Phi(j_2,\gamma)\right)$;
10. Sphericity [11];
11. $\max\left(\Delta\eta(j_1,\gamma), \Delta\eta(j_2,\gamma)\right)$;
12. $\min\left(\Delta\eta(j_1,\gamma), \Delta\eta(j_2,\gamma)\right)$;
13. $\Delta\Phi(\gamma, W)$ where W is the jet1-jet2 system;
14. η of the second jet;
15. $E_T^{jet1} - E_T^{jet2}$;
16. β of the jet1-jet2 system;
17. Aplanarity[11];
18. $\cos\theta^*(j_1, \gamma)$ where the angle is calculated in the $j_1 - \gamma$ rest frame;
19. $\Delta\eta(\gamma, W)$;

[10] The counting is made at the vertex

[11] The sphericity tensor is defined:

$$S^{\alpha\beta} = \frac{\sum_i p_i^{\alpha} p_i^{\beta}}{\sum_i |\mathbf{p}_i|^2}$$

where i is the jet index, α, β=1,2,3 correspond to the x, y and z components. By standard diagonalization of $S^{\alpha\beta}$ one may find three eigenvalues $Q_1 \geq Q_2 \geq Q_3$, with $Q_1 + Q_2 + Q_3 = 1$. The sphericity is then defined as $S = \frac{3}{2}(Q_2 + Q_3)$, so that $0 \leq S \leq 1$.
The aplanarity is defined as $A = \frac{3}{2}Q_3$, so that $0 \leq A \leq 0.5$.

By comparing the signal and the background distributions of each of these 19 variables we found that they all have some discriminatory power. However, it is not hard to predict that some of these quantities are correlated with each other. Adding correlated variables to the network clearly does not improve the signal/background separation. The information they bring in to the network have already been exploited using the other (correlated) variables. As a general rule it is always recommended to use as inputs only those variables that maintain some degree of discrimination power even when they are *combined* with the other inputs. One might think that this step is not really necessary since an overabundant set of input variables would not affect the performance of the network or its final outcome. After all, the neural network will spot these correlations during the training and simply will ignore the information of the correlated variables. Unfortunately this is true only in principle. As we have explained in the previous section, the minimization of the error surface is not done analytically but it is a sort of "empirical" exploration in search for the global minimum. Increasing the dimensions and the complexity of the surface does not make the process any easier. Therefore it is always advised to build a NN with as few degree of freedom as possible to help the stability convergence during the network training.

The Short List Input Variables

In order to reduce the variable input list and retain only a minimum set of them with the *same* (or very similar) background rejection power, we performed a study using the following procedure:

- Train a NN with each variable separately, and rank them according to their performance from the best to the worse.
- Train a 2-dimension NN with the best variable found before and one of the other variables (one at time), and identify the best pair of variables.
- (*After k steps*) Training a k-dimensions NN [12] , add one variable at time to the best k-1 variables found in the previous step and identify the best k variables.
- Repeat the procedure until all variables are included in the best-variable set.

We basically formed all the possible combinations of the input variables [13] and we calculated their performances defined as the $S/\sqrt{B}$ after the cut on the NN output that keeps 75% of the signal. The best combinations are thus those with the higher significance. What we expect from this study is to see the best NN selection significance increasing with the number of input variables. In fact after every step more information becomes available to carry out a better

[12] The number of hidden neurons was $n_h = k + 2$

[13] To do this study we needed to train $N \cdot (N+1)/2 = 19 \cdot (19+1)/2 = 190$ NN!

discrimination. This trend will last until the uncorrelated variables run out. At this point adding more variables will not improve further the significance because they are correlated with variables already included in the network. In this way we will know both how many *uncorrelated* inputs (out of the 19 we started with) we have and which are they. In fig. 9.4 we plot as a function of the number of inputs k, the best and worse performance among all the possible input combinations (for a given k) [14]. We can see the feature we have anticipated.

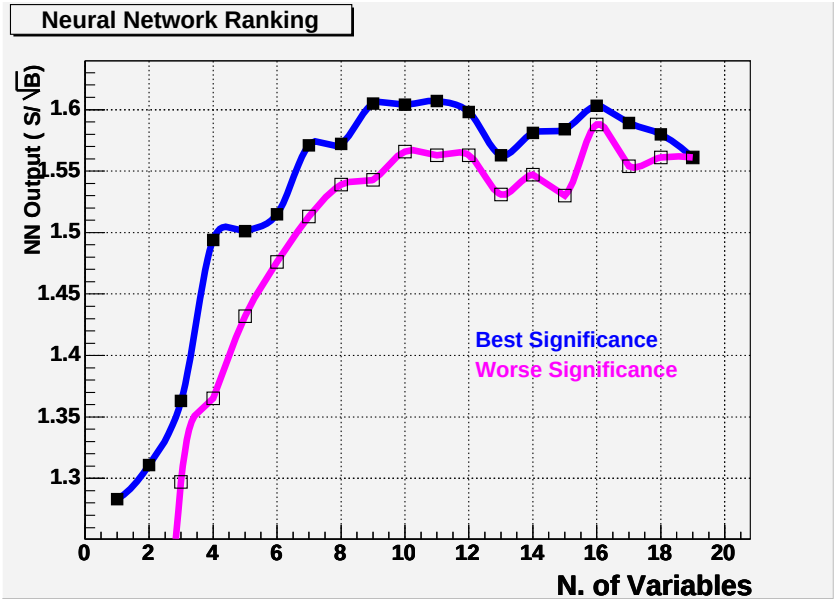


Figure 9.4: *Best (blue) and worse (magenta) NN input variable combination in term of significance as a function of the number of input nodes.*

The significance increases until a sort of plateau is reached (around N=10) at which point the performance of the network is basically constant in spite of the increasing number of input variables. The significance values for all the 190 combinations are inside the region between the two curves shown in the plot whose size can be interpreted as an error band [15].
As a results of this study we narrowed the number of inputs from 19 to 10 and these selected ones are the first ten reported in the input list above [16].
These ten variables are the input nodes for the NN we used for our advanced

[14] For k input variables there are 20-k combination.

[15] The fluctuation of the curves are due to little changes in the NN parameters we need to introduce when the number of inputs become large. These parameters have only a little impact on the training performance.

[16] This didn't happen by chance. Indeed the variable order in the list was not random but we have sorted the variables out according to the study we have just described, but, of course, we couldn't mention that before introducing the ranking method itself!

selection. The distributions of these variables in the data and in the signal Monte Carlo are compared in fig. 9.5- 9.6.

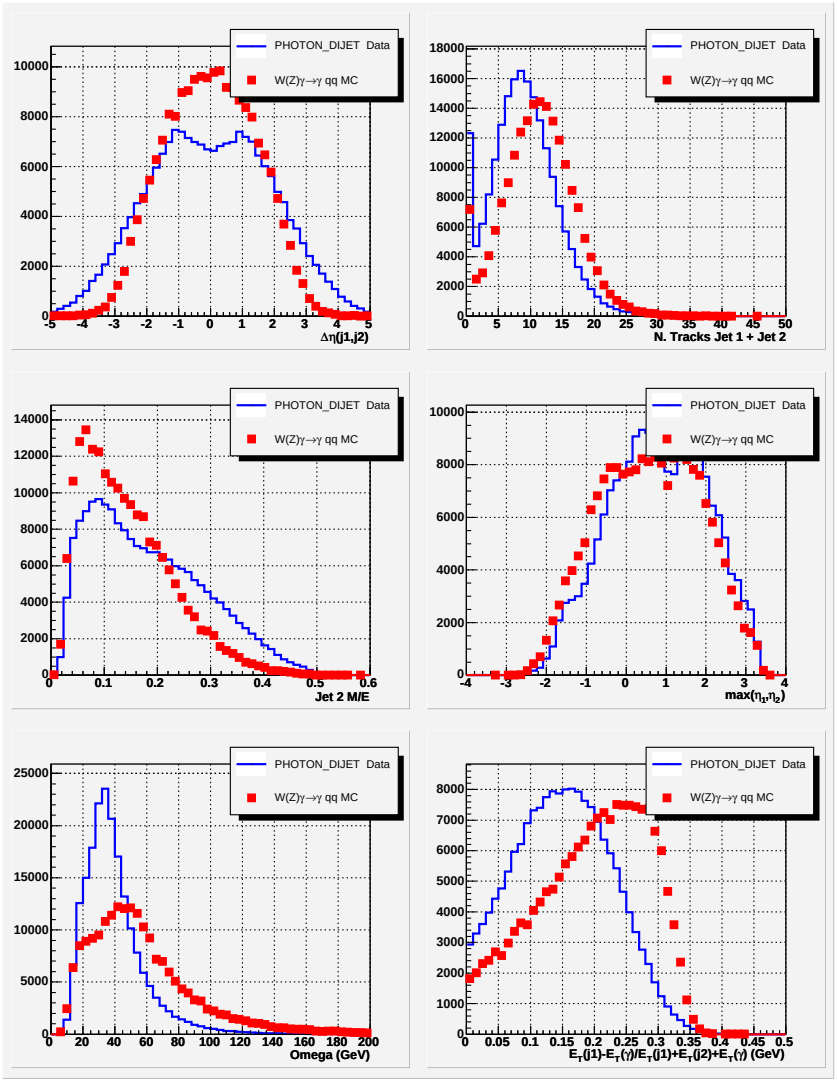


Figure 9.5: *Distributions of the first six input variables for the signal Monte Carlo (full line) and the data (red points). From the left to the right, top to the bottom we have: 1)* $\Delta\eta(J_1, J_2)$ *; 2) Number of tracks in the first two jets; 3)* M/E *of the second jet; 4)* $\max(\eta^{jet1}, \eta^{jet2})$*; 5)* Ω *; 6)* $(E_T^{jet1} - E_T^{\gamma})/(E_T^{jet1} + E_T^{jet2} + E_T^{\gamma})$

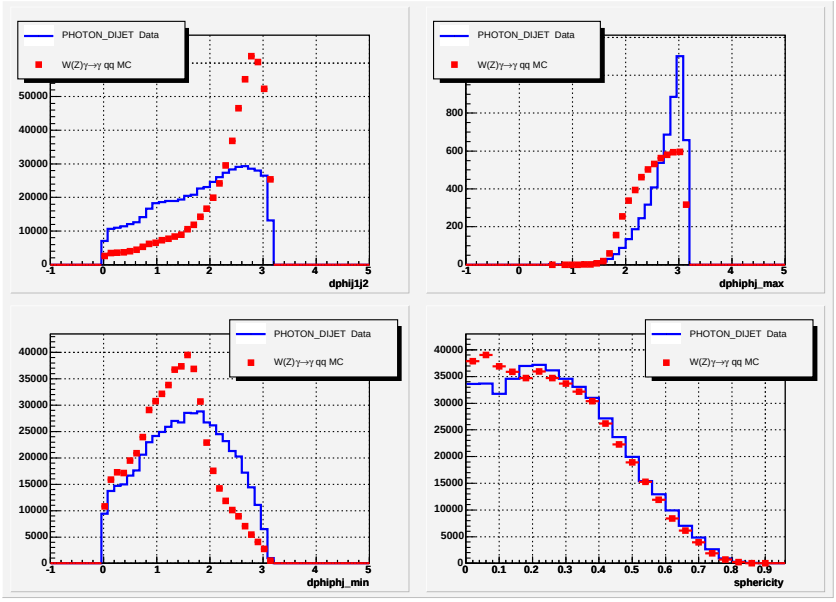


Figure 9.6: *Distributions of the last four input variables for the signal Monte Carlo (full line) and the data (red points). From the left to the right, top to the bottom we have: 7)*$\Delta\Phi(j_1, j_2)$ *; 8)* $\Delta\Phi(j_{MAX}, \gamma)$ *; 9)* $\Delta\Phi(j_{MIN}, \gamma)$ *; 10) Sphericity ;*

9.3.3 Network Training

In order to complete the construction of our NN we have to decide the number of hidden nodes we want to use. To optimize this choice we studied the significance of the NN selection as a function of the hidden variables in a similar way we did for the input variables. Spanning the hidden nodes from 11 to 30, we recorded only mild differences in significance and we didn't notice any particular trend. This is in agreement with the observation of a weak dependence of the network performance upon the number of hidden nodes reported in other analysis [50]. As a result we fixed - rather arbitrary - the $N_h = 17$.
It is a widely accepted rule to use for a proper training a number of events which is 20-40 times the number of degree of freedom of the NN, given by the following formula:

$$N_{dofF} = N_h \cdot (N_I + N_O + 1) + N_O$$

Since in our case we have $N_{dofF} = 205$ we used 8000 events for the training [17] As the learning rate and momentum parameters of the back-propagation learning algorithm is concerned we choose the default values. The number of *epochs* was set to 10000.

[17] As a reference we recall that we have 40K events in the Monte Carlo sample and about 190K in the data (90K inside the signal region) after all the standard selection cuts.

After having decided all the parameters of the network we prepared the two training samples - one from the signal and the other for the background - with 8K events each randomly selected from the Monte Carlo $W(Z)\gamma \to \gamma q\bar{q}$ sample and from the data sample. It is important to stress that for the background template we peaked only data events *inside* the dijet mass signal region. This decision was prompted by the fact that we do not want the NN looking for the differences between the data *outside* the W/Z peak mass window and the signal because (1) they are useless as far as the discrimination in the signal region is concerned and (2) there is a risk to sculpt the dijet mass distribution or move its turn-on.
With the two training templates we finally constructed our NN and its performance on our dataset is discussed next.

9.4 Neural Network Output

Once the training is completed, all the synaptic weights and thresholds of all the neurons in the network are determined. Now we can apply the neural network to any event we want and it will output a number (N_{OUT}) between 0 and 1 [18]. The first thing we did is to apply the NN to both our signal and data sample. Comparing the two N_{OUT} distributions (fig. 9.7) one can appreciate the stunning discrimination power of our NN. We should bear in mind that all these events have already been filtered with the standard selection cuts. But - as it should be clear now - they were not able to take a full advantage of the several differences between the signal and the background, in contrast to what the NN selection does. In order to choose the best cut to apply on N_{OUT} we plotted both the efficiency of the signal (ϵ_S) and of the background (ϵ_B) as a function of N_{OUT} (fig. 9.8) and the efficiency of the signal versus the signal over background ratio (fig. 9.9). This second plot is particularly interesting because shows that two desiderata - maximum S/B and minimum signal loss - are in conflict. To make up our mind we plot also the "relative" significance (defined as $\epsilon_S/\sqrt{\epsilon_B}$) as a function of N_{OUT} (fig. 9.10). It shows that a cut at 0.75 would give the highest significance. Unfortunately, for this cut the corresponding ϵ_S is only 50%, an efficiency that we gauged too low to be acceptable.
In conclusion, as a result of the study made with the neural network, we decided to apply an extra selection criteria on our dataset - besides the standard cuts described in the previous Ch.7 - requiring the events to have the output of the NN $N_{OUT} > 0.6$.

9.5 Consistency Check and Selection Efficiency

It is very important, before going any further, to check the shape of the dijet mass distribution *after* the NN cut and in particular if the position of the dijet mass turn-on has been shifted as a result of the new selection criteria. In fig. 9.11

[18] We recall that 0 was the target output for the background and 1 was for the signal.

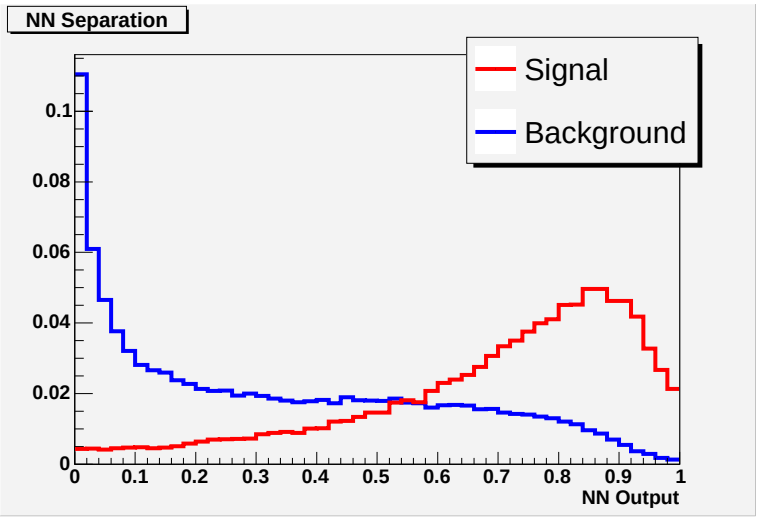


Figure 9.7: *The N_{OUT} distributions for the signal and for the background events.*

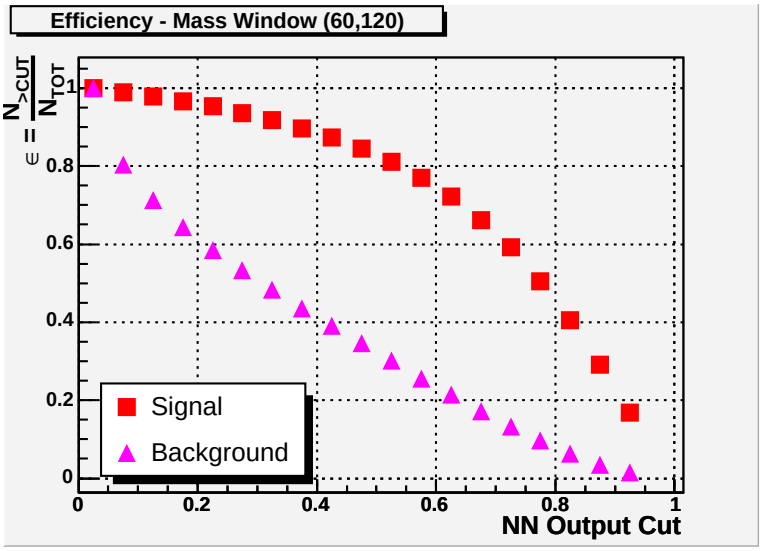


Figure 9.8: *Efficiency for the signal and the background as a function of the NN output N_{OUT}.*

the dijet mass spectra for different N_{OUT} are shown. We can notice that not only the turn-on did not move toward high values but it is even lower than it was without the NN cut. As sanity check we looked as well at the signal peak

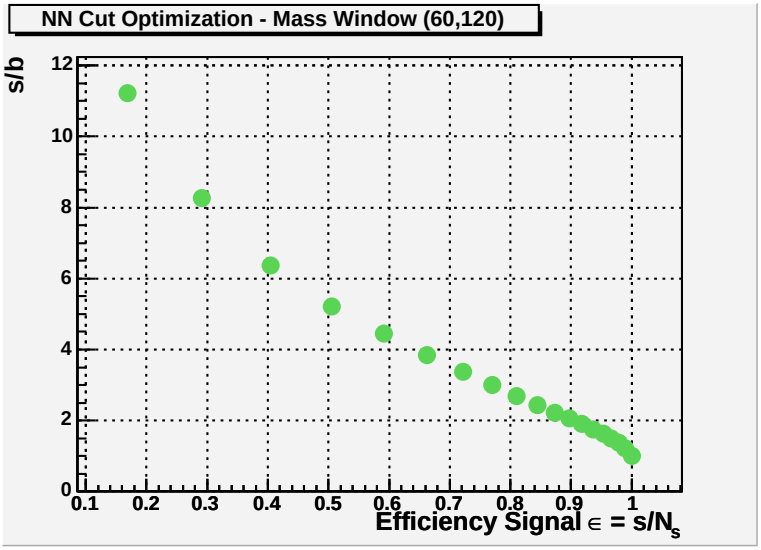


Figure 9.9: *Signal over Background efficiency ratio (ϵ_S/ϵ_B) as a function of the signal efficiency ϵ_S. This plot has been obtained from the $\epsilon_B(N_{OUT})$ and $\epsilon_B(N_{OUT})$ reported in fig. 9.8 and eliminating N_{OUT}. This plot shows how an high signal efficiency is in conflict with an high S/B.*

distributions (fig. 9.12). As expected the effect of the NN cut is very small and it seems also that the dijet mass resolution even improves a little bit.

Multiple Interaction Effects

Another important point to be checked is the dependence of our NN upon the amount of multiple interactions (MI). As a matter of fact almost half of the events in the data sample have more than one $p\bar{p}$ interaction with an average number of vertices of 1.7 [19]. On the other hand the Monte Carlo sample does not have any extra $p\bar{p}$ collisions. A reason of concern is that this intrinsic difference between the two NN templates could be spotted by the NN and interpreted as a real physics effect, and therefore be used improperly in the discrimination decision. In other words, the shape of the input variables sensitive to MI can be distorted in the data sample, causing a bias in the NN selection.
The first thing to notice is that most of our inputs include only the two leading jets and the photon information and they are not expected to be sensitive to the MI (we will verify that in a moment). In fact the jet and photon directions are not affected by MI and their energies are corrected for this effect (sec. 4.6.2).

[19] The average numbers of interaction depends upon the instantaneous luminosity and as a consequence on the data taking period. For example the `PHOTON_DIJET_L2_DPS` dataset has a lower average (1.4) because it operated only until January 2004.

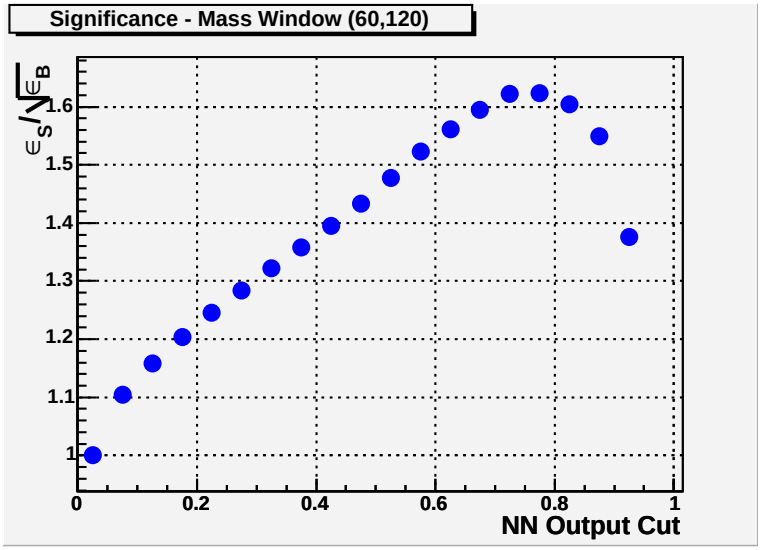


Figure 9.10: *NN Significance (defined as $\epsilon_S/\sqrt{\epsilon_B}$) as a function of N_{OUT}. This plot is the best way to decide the optimum N_{OUT} cut.*

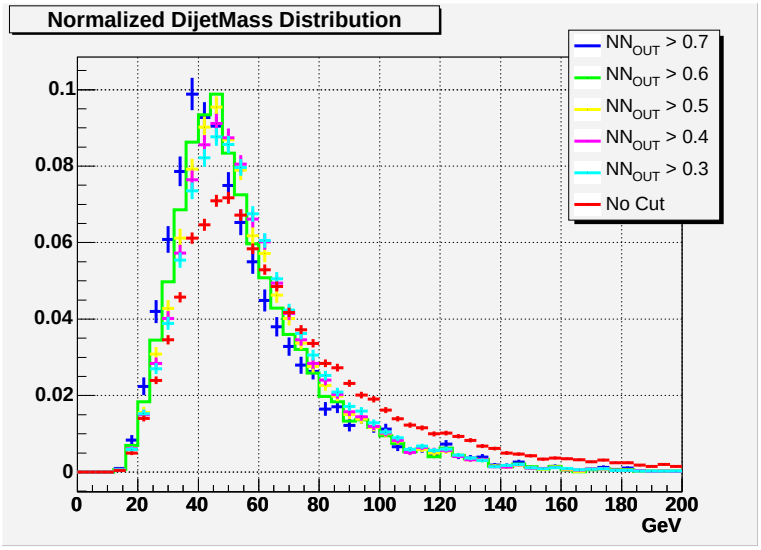


Figure 9.11: *Normalized dijet mass distribution for the data after the NN selection. Several different cuts on N_{OUT} are shown. All of them do not have any effect on the position of the turn-on.*

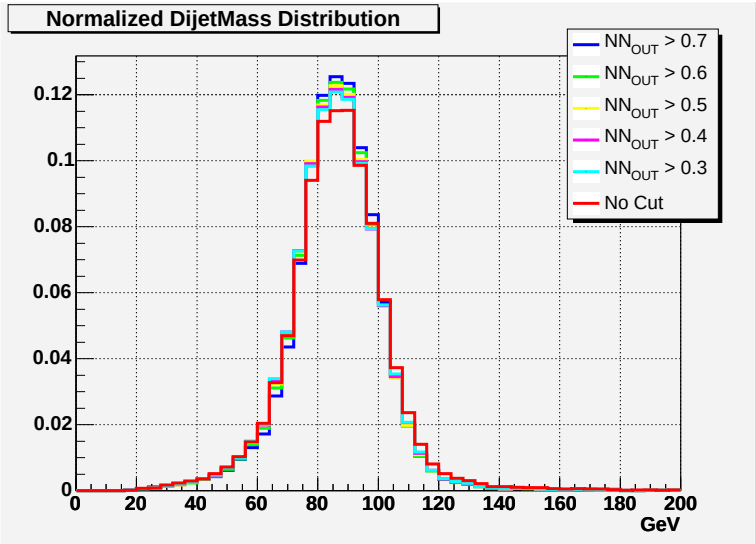


Figure 9.12: *Normalized dijet mass distribution for the signal after the NN selection. Several different cuts on N_{OUT} are shown. There is not any appreciable change in shape except maybe a tiny improvement in the dijet mass resolution.*

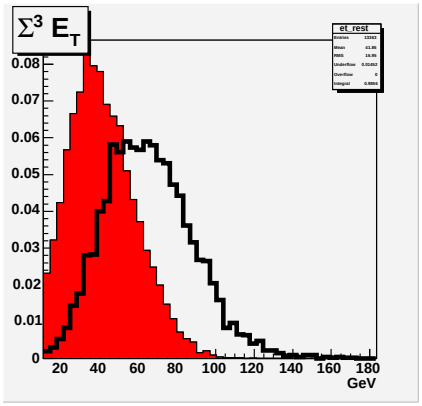
Moreover we never use absolute values but only differences of energies, canceling out possible dependences.
There is one variable though that entails a more careful study: that one we called the "intrajet energy" Ω (the 5^{th} in the input list). From its definition it is clear that the use of the unclustered tower energies could in principle lead to some dependence on the extra energy coming from the MI. The introduction of this quantity was prompted by the color flow differences between involved diagrams of the signal and the background. In fact, the signal is characterized by the presence of two colorless objects that constrains the initial and final state in a particular (singlet) color configuration. The QCD background is instead made of several different possible configurations (see fig. 8.1) involving quarks and gluons in the initial and final state with multiple color connections. This implies higher color charge and higher color radiation with, moreover, a different pattern with respect to the signal diboson case.
To take advantage of all these differences one needs to find an observable that is sensitive to them. For what we just said, this variable should involve the total amount of color radiation. One of the possibilities is to use the total transverse energy in the event ($\sum E_T$) *excluding* the three leading objects, the photon and the first two jets (*i.e.* $\sum^{>3} E_T = \sum E_T - E_T^{jet1} - E_T^{jet2} - E_T^{\gamma}$ [20]). However this

[20] A similar quantity is used in other analyses involving jets in the final state coming from vector boson decays [51]. They actually count only the transverse energy of the reconstructed jets in the events instead of the total transverse energy.

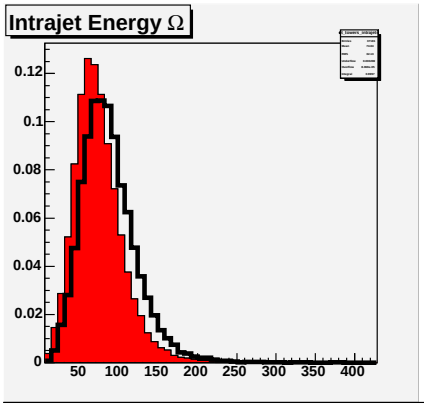
quantity can be sensitive to the energy from the MI. To check that, we divided our data sample in two subsets, the first made of events with only one vertex, the second made of events with more than one. Then we compared the $\sum^{>3} E_T$ distributions in the two subsamples (fig. 9.13(a)). It is clear that this quantity strongly depends on the MI and "as it is" cannot be used to train the NN. We therefore started to modify its definition and we ended up with the quantity Ω. As it can be seen in fig. 9.13(b) Ω does not suffer the MI dependence and it is still sensitive to the color radiation effects.
Considering the importance of the MI contamination we made another study to verify that *all* the variables we use in the NN are not affected by the MI. For this reason we built a dedicated NN trained with the two aforementioned subsamples, one used as a signal template and the other as background. As a consequence, such a network was trained to discriminate events with one vertex from those with more than one. Of course this NN can successfully do that only if it has been provided of input variables sensitive to the number of vertices. Otherwise, with the information at its disposal, it will fail the discrimination. In reality this is exactly what happened, as shown in fig. 9.14, and this fact can be interpreted as a strong evidence that our variable set is not biased by the presence of extra $p\bar{p}$ collisions.

NN Selection Efficiency

As we did with all the other selection criteria, we need now to evaluate the NN cut efficiency ϵ_{NN} and its uncertainty. Since the NN selection is applied on top of the other cuts, the relative efficiency is just the fraction of signal events with $N_{OUT} > 0.6$. The result is $\epsilon_{NN} = 0.70$.



(a) $\sum^{>3} E_T$



(b) Ω

Figure 9.13: *Comparison of the (a) $\sum^{>3} E_T$ and (b) Ω distribution for events with 1 vertex only (full red) and events with two or more (empty black).*

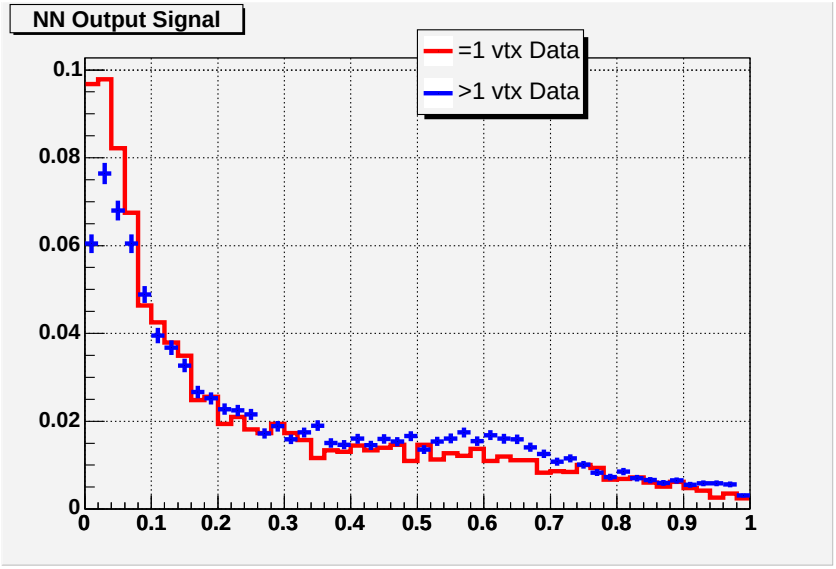


Figure 9.14: *Output distribution of a NN trained to separate events with only one vertex from events with more than one. Using the input variables we employ in our analysis NN, it is not able to separate them proving that our set of inputs are not sensitive to the MI.*

To estimate the uncertainty we first notice that the input variables of the NN are all but two related only with the energy and the position of the photon and the two leading jets. As a consequence their errors involve only the jet energy scale uncertainties. With the same procedure adopted to study the jet cuts systematics, we can shift by 1σ the jet energy scale and set the efficiency differences as the ϵ_{NN} systematic uncertainty (in this case we treated the photon as a jet of a 0.4 cone radius). For two variables - Ω and the number of tracks (n.2 in the input variable list) - we need instead a different approach. The intrajet energy Ω depends upon the calorimeter tower energy [21] and we cannot use the 1σ scale shift used before because there is no corrected jet here. What we need is to evaluate the systematic error on the Ω Monte Carlo prediction. With good approximation we can assume that the same systematic error affects the $\gamma+2jet$ `Madgraph` Monte Carlo events (sec. 8.2) being the error it is correlated with the detector simulation, common to all the Monte Carlo samples. Hence we can compare the Ω distributions of the $\gamma + 2jet$ Monte Carlo sample with the data and use their difference as uncertainties for Ω (fig. 9.15). The fact that the $\gamma + 2jet$ Monte Carlo sample is not able to reproduce the data (as we discussed in sec. 8.3) does not void our estimate but, at most, makes it more conservative. A similar approach was followed for the systematics on the other variable, the

[21] In the Ω definition the jet and photon energy are *not* corrected, *i.e.* only the calorimeter tower energy in the cone are used.

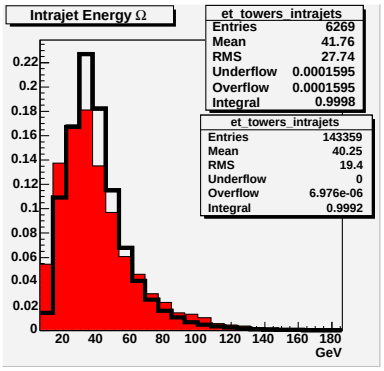


Figure 9.15: *Intrajet energy Ω distribution for the $\gamma + 2jet$ Monte Carlo sample (full red) and for the data (empty black). A difference of 3.7% is observed and used as a systematic uncertainty.*

number of tracks associated with the two leading jets.
One important fact we need to take into account is the correlation between the cut efficiency on the jet quantities (ϵ_{jets} and ϵ_{NN}). Since they all depend, in a way or another, from the jet energy scale errors, we cannot add them in quadrature but instead they need to be calculated together. Starting from the events that passed just the photon selection, we determined the *final* efficiency (with the NN selection included) changing by $\pm 1\sigma$ *both* the jet energy scale, Ω and the number of tracks in the first two jets, all at the same time. In this way we set a 27% systematics on $\epsilon_{jets} \cdot \epsilon_{NN}$.

9.6 Mass Window Optimization

After the control checks and the efficiency study, it is time now to calculate the new significance after the NN selection. Computing the number of events in the signal region 60-120 GeV we found a S/B=1/71 and a significance $\frac{S}{\sqrt{S+B}} = 1.51$, a big improvement with respect to the S/B=1/333 and $\frac{S}{\sqrt{S+B}} = 0.80$ obtained before applying the NN advanced selection.
Moreover, it has to be noted that since the signal distribution is peaked at some certain mass, the wide dijet mass window used to compute the number of events is not the optimal choice. To identify the best dijet mass window we varied the extremes of the dijet mass interval in order to maximize the significance. We found the interval 72-110 GeV (fig. 9.16) to be our best mass window , corresponding to a S/B=1/41 and $\frac{S}{\sqrt{S+B}} = 1.86$.

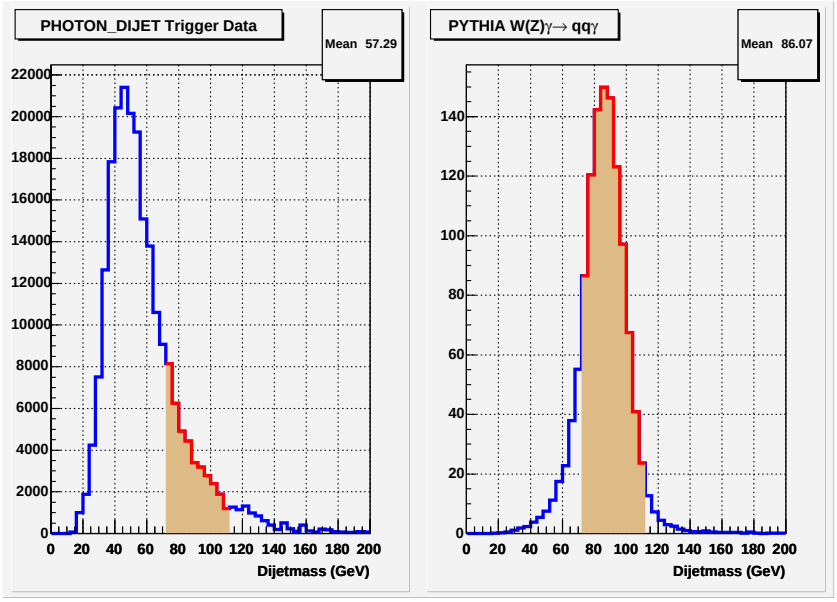


Figure 9.16: *Dijet Mass spectrum for the background and signal* after *the NN selection cut. The highlighted region is optimized to maximize the* $\frac{S}{\sqrt{S+B}}$.

9.7 Background Fit and Peak Search

After all the work selecting and optimizing the best possible photon dataset for the search of the hadronic decays of the W and Z, we are finally arrived at the crucial point of this analysis. We will now fit the two sidebands of the dijet mass distribution and we will determine the shape of the background extrapolating them inside the signal region. Then we will compare the data points with the expected background and see if there is an excess of events that can be attributed to the W/Z resonance production.
First of all we have to decide how to fit the background. We tried three possible functions and we realized that the exponential was by far the best choice in term of fit stability [22] (fig. 9.17). Since we have two sidebands to fit, we need to fix their size. We define the low mass sideband as the region $[M_{ll}, M^{jj}_{min}]$ and the other sideband as $[M^{jj}_{max}, M_{rr}]$. The signal region is of course the interval in between $[M^{jj}_{min}, M^{jj}_{max}]$. [23]. Finally we have to set the distribution bin size ΔM.

We performed several fits varying all these parameters and we found an ex-

[22] The other two functions we tried were: (1) $x^{-a}e^{-bx}e^{-cx^2}$ and (2) $x^{-a}\ln\frac{b}{x}\ln\frac{b}{x^2}$

[23] The signal region used for the peak search can be different in principle to the optimized mass window we have found maximizing S/B. The main concern here is the stability of the fit and the accuracy in the determination of the background shape.

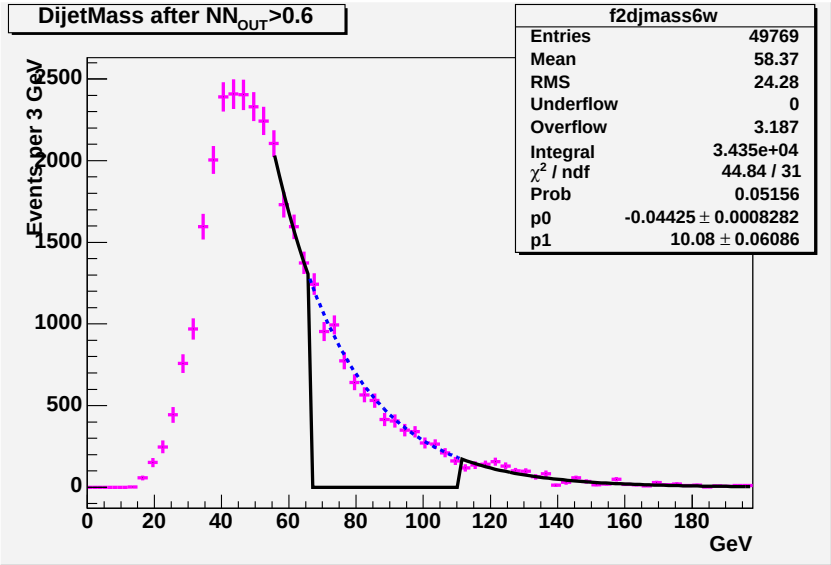


Figure 9.17: *Dijet Mass distribution of the data after the NN selection cut. An exponential function $e^{p_0+p_1\cdot x}$ is used to fit the two sidebands (full line) and the result is extrapolated inside the signal region (dashed line).*

traordinary stability of the fit outcomes. This nice behavior has to be attributed to the presence of two segments in the fit that constrains it very strictly. All the struggling in the construction of the trigger and in the event selection to keep a low mass turn-on eventually paid off.
We summarized the fit results in three tables where the changes in ΔM (Table 9.1), in $[M^{jj}_{min}, M^{jj}_{max}]$ (Table 9.2) and in M_{ll} (Table 9.3) are reported. As an example, the effect of different bin size reported in Table 9.1 are shown in fig. 9.18.

As we can see, except for few cases where `MINUIT`, the minimization program, does not converge, the fit results are very similar for all possible configurations. Therefore the setting of the fit parameters is almost an arbitrary choice.
For the peak search we set the bin size to 4 GeV, the lower limit of the low sideband M_{ll} to 52 GeV and the mass window to $[68-116]$ GeV. But we have to bear in mind that other choices would point to the same results.
In fig. 9.19 we show the corresponding dijet mass spectrum along with the background fit and in fig. 9.20 we subtracted the background contribution from the data to seek for a possible excess.
As we somehow suspected, in spite of all the optimization we have introduced, the statistics of the sample is still too poor for the W/Z peak to emerge.
In the next chapter we will set a limit on the $W(Z)\gamma \to \gamma q\bar{q}$ cross section and we will give an estimate of the future achievable sensitivity in Run II.

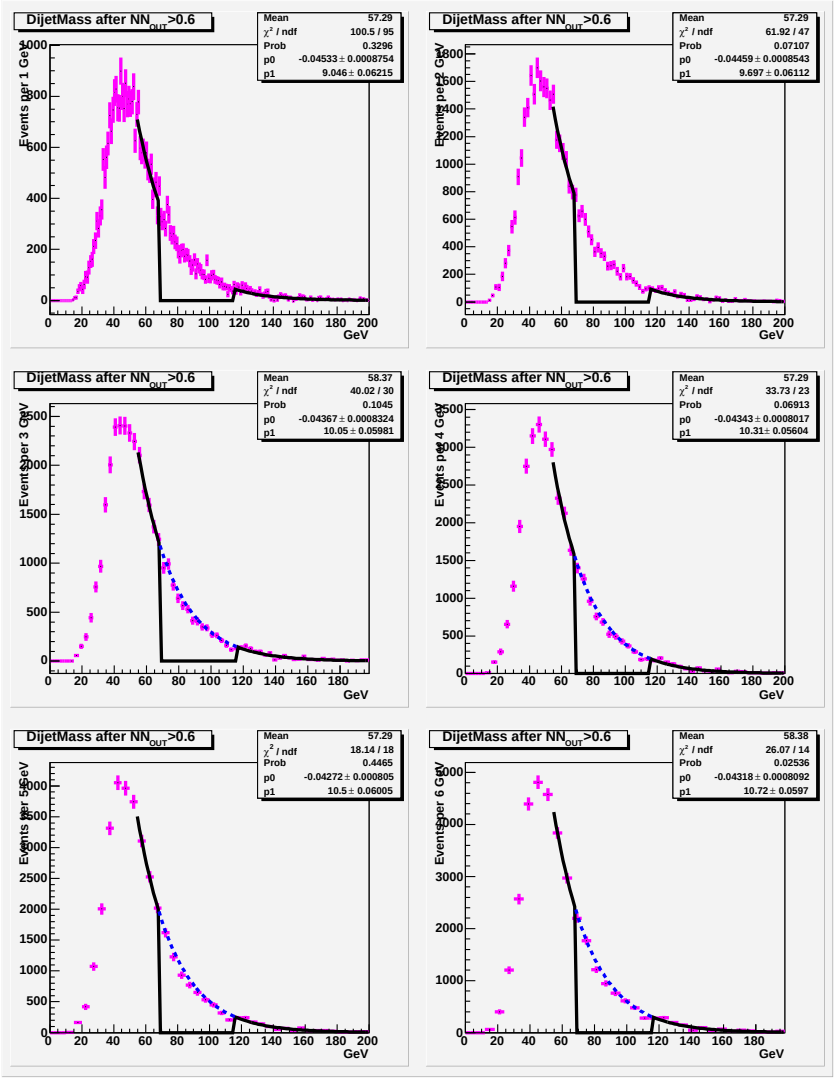


Figure 9.18: *Exponential fit of the sidebands of the data dijet mass distribution for different histogram bin size. The results of the fits are reported in Table 9.1.*

Fit Parameters for Different ΔM					
ΔM (GeV)	M_{ll}	$[M^{jj}_{max}, M^{jj}_{max}]$	p_0	p_1	χ^2/ndf
1.0	55	72-110	-0.0457±0.0009	9.1± 0.06	114.1
2.0	55	72-110	-0.0451±0.0008	9.7± 0.06	68.2
3.0	55	72-110	-0.0443±0.0008	10.1± 0.06	48.1
4.0	55	72-110	-0.0189±1.0106	7.6± 1.00	1750.9
5.0	55	72-110	-0.0435±0.0008	10.5± 0.06	26.0
6.1	55	72-110	-0.0439±0.0008	10.7± 0.06	33.0
7.1	55	72-110	-0.0427±0.0008	10.8± 0.06	12.2
8.0	55	72-110	-0.0428±0.0008	11.0± 0.06	17.5

Table 9.1: *Sidebands fit results when different bin size ΔM of the data dijet mass histogram are adopted.*

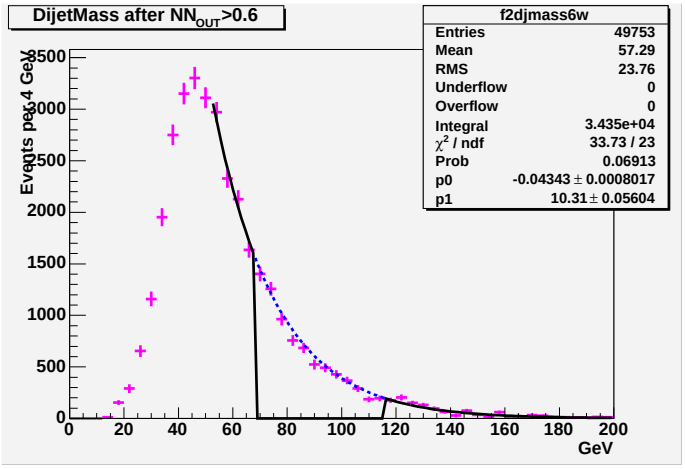


Figure 9.19: *Dijet Mass distribution of the data after the NN selection cut. An exponential function $e^{p_0+p_1*x}$ is used to fit the two sidebands (full line) and the result is extrapolated inside the signal region (dashed line). The parameter for the fit are those used for the peak search: ΔM=4 GeV; M_{ll}=52 GeV; $[M^{jj}_{max}, M^{jj}_{max}] = [68-116]$ GeV.*

Fit Parameters for Different Mass Window					
ΔM	M_{ll}	$[M^{jj}_{max}, M^{jj}_{max}]$	p_0	p_1	χ^2/ndf
3.0	54	60-111	-0.0442±0.0008	10.1± 0.07	44.7
3.0	54	60-114	-0.0438±0.0008	10.1± 0.07	40.1
3.0	54	60-117	-0.0437±0.0008	10.0± 0.07	39.8
3.0	54	60-120	-0.0438±0.0009	10.1± 0.07	39.6
3.0	54	60-123	-0.0267±1.0462	8.4± 1.00	762.5
3.0	54	60-126	-0.0445±0.0010	10.1± 0.07	34.1
3.0	54	63-111	-0.0443±0.0008	10.1± 0.06	44.8
3.0	54	63-114	-0.0438±0.0008	10.1± 0.06	40.1
3.0	54	63-117	-0.0207±1.0073	7.0± 1.00	1223.9
3.0	54	63-120	-0.0438±0.0009	10.1± 0.06	39.6
3.0	54	63-123	-0.0441±0.0009	10.1± 0.07	36.5
3.0	54	63-126	-0.0445±0.0010	10.1± 0.07	34.2
3.0	54	66-111	-0.0443±0.0008	10.1± 0.06	44.8
3.0	54	66-114	-0.0438±0.0008	10.1± 0.06	40.2
3.0	54	66-117	-0.0279±1.0021	8.6± 1.00	824.3
3.0	54	66-120	-0.0438±0.0009	10.1± 0.06	39.6
3.0	54	66-123	-0.0441±0.0009	10.1± 0.06	36.5
3.0	54	66-126	-0.0445±0.0010	10.1± 0.07	34.2
3.0	54	69-111	-0.0443±0.0008	10.1± 0.06	45.1
3.0	54	69-114	-0.0438±0.0008	10.1± 0.06	40.3
3.0	54	69-117	-0.0437±0.0008	10.1± 0.06	40.0
3.0	54	69-120	-0.0438±0.0009	10.1± 0.06	39.8
3.0	54	69-123	-0.0441±0.0009	10.1± 0.06	36.8
3.0	54	69-126	-0.0445±0.0010	10.1± 0.07	34.5
3.0	54	72-111	-0.0443±0.0008	10.1± 0.06	48.1
3.0	54	72-114	-0.0438±0.0008	10.1± 0.06	43.7
3.0	54	72-117	-0.0437±0.0009	10.0± 0.06	43.5
3.0	54	72-120	-0.0438±0.0009	10.1± 0.06	43.2
3.0	54	72-123	-0.0442±0.0009	10.1± 0.07	39.8
3.0	54	72-126	-0.0446±0.0010	10.1± 0.07	37.3

Table 9.2: *Sidebands fit results when different signal mass window definitions* $[M^{jj}_{max}, M^{jj}_{max}]$ *are adopted.*

Fit Parameters for Different M_{ll}					
ΔM	M_{ll}	$[M^{jj}_{max}, M^{jj}_{max}]$	p_0	p_1	χ^2/ndf
3.0	51	72-111	-0.0177±1.0049	7.2± 1.00	2687.4
3.0	54	72-111	-0.0443±0.0008	10.1± 0.06	48.1
3.0	57	72-111	-0.0440±0.0009	10.1± 0.07	47.2
3.0	60	72-111	-0.0442±0.0009	10.1± 0.07	46.9
3.0	63	72-111	-0.0440±0.0010	10.0± 0.08	46.6
3.0	66	72-111	-0.0191±1.0671	7.0± 1.00	618.0
3.0	69	72-111	-0.0427±0.0013	9.9± 0.14	43.5

Table 9.3: *Sidebands fit results for different lower edge definition (M_{ll}) of the left sidebands are adopted.*

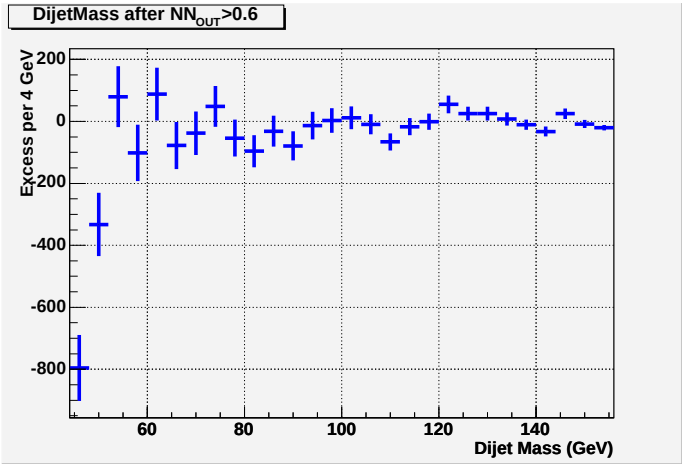


Figure 9.20: *Excess of events in the data with respect to the background prediction deduced from the sideband fits. The turn-on effect can be noticed in the first two bins. No evidence of an excess from the W/Z resonance production is found inside the signal region.*

Chapter 10

Upper Limit and Future Prospects

10.1 Introduction

At the end of the last chapter we showed how in 184 pb^{-1} of data collected with the new dedicated PHOTON_DIJET trigger the overwhelming QCD background present in the sample did not allow us to identify the W/Z peak signal, despite the use of an optimized advanced selection.
Given the negative result of the search, in this chapter we will set an upper limit on $\sigma(p\bar{p} \to W/Z\gamma) \times \mathfrak{B}(W/Z \to q\bar{q})$. We will start (sec. 10.2) with the calculation of the limit by assuming to know exactly the number of expected signal and background events (*i.e.* neglecting the uncertainty), and taking the opportunity to explain in detail the method employed. Then (sec. 10.3) we will generalize the study showing how to fold the systematic uncertainties in the calculation of the limit. Our computations - under whatever assumption - will always be based on a fully Bayesian approach.

10.2 Upper Limit Calculation

In order to illustrate the method we are going to use, we start analyzing a very simple scenario of an upper limit calculation for a process with an expected cross section σ. Let's suppose we performed a counting experiment observing n events. Among them we expect N_b events from the background and $N_s = \epsilon\sigma L$ events from the signal, where ϵ is the total acceptance and L is the integrated luminosity of the sample. Let's assume also that we know all these numbers *exactly*. The probability to find n counts when we expect $\mu = \epsilon\sigma L + N_b$ is given by:

$$\mathcal{L}(\mu|n) = \mathcal{L}(\sigma, \epsilon, L|n) = \frac{\mu^n}{n!}e^{-\mu}$$

In Bayesian statistics, the information about the parameters σ, ϵ and L can be represented by a probability distribution. Therefore a *prior* density $\pi(\sigma, \epsilon, L)$ can be defined which contains the knowledge we have about these parameters *before* any measurement. The fact that we actually count n events modifies our knowledge about these parameters and Bayes theorem provides us with the updated, *posterior* density:

$$p(\sigma, \epsilon, L|n) = \frac{\mathcal{L}(\sigma, \epsilon, L|n)\pi(\sigma, \epsilon, L)}{\iiint \mathcal{L}(\sigma', \epsilon', L'|n)\pi(\sigma', \epsilon', L')\ d\sigma'\ d\epsilon'\ dL'}$$

where the denominator is a posterior normalization factor that does not depend upon the parameters σ, ϵ and L but only on n (in this way, integrating $p(\sigma, \epsilon, L|n)$ over all the parameter space gives the right normalization, 1).
In reality, since we want to calculate a limit on the cross section, we are interested only in the posterior density for σ. The latter can be obtained by integrating the joint posterior over other two parameters:

$$p(\sigma|n) = \iint p(\sigma, \epsilon', L'|n)\ d\epsilon'\ dL'$$

At this point to calculate the 95% confidence level (C.L.) upper limit σ_{95} we have just to solve this equation

$$\int_0^{\sigma_{95}} p(\sigma|n)\ d\sigma = 0.95$$

and find the unknown σ_{95}.
There is still one thing, however, that we need to do to solve the problem: define the *prior* density $\pi(\sigma, \epsilon, L)$. Since the three parameters are independent, we can separate them and write $\pi(\sigma, \epsilon, L) = \pi(\sigma)\pi(\epsilon)\pi(L)$ [1]. Moreover, because in this example we assumed a perfect knowledge of ϵ and L, their prior densities are clearly Dirac delta functions. For the cross section instead, we assume a uniform prior, $\pi(\sigma) = 1$. It is important to note however that in general it is not advised to adopt a flat prior because it may lead to a posterior density whose normalization diverges (see [52]).
Given the prior density we can now calculate the posterior density for σ as:

$$p(\sigma|n) = \frac{\mathcal{L}(\sigma, \epsilon, L|n)}{\int_0^\infty \mathcal{L}(\sigma', \epsilon, L|n)\ d\sigma'}$$

[1] The independence is trivial in this example, but we will assume it also in the more realistic case we will discuss later on. The measurement of L (and its error) is definitely independent from the σ and the efficiency of our analysis. The expected σ is also independent from the efficiency since the Monte Carlo generator - that calculated the cross section - doesn't know about the selection cuts we were going to apply. The contribution to the efficiency coming from the detector simulation can be also considered independent. It is not clear however if the efficiency on the $P_T^\gamma > 12$ GeV cut is independent from the cross section calculation. A different PDF or factorization scale may change the photon energy distribution altering the fraction of events above $P_T^\gamma > 12$ GeV. We did not investigate further this issue because we gauged this effect much smaller than the systematic uncertainty we already have.

and the equation to solve becomes

$$\frac{\int_0^{\sigma_{95}} \mathcal{L}(\sigma', \epsilon, L|n)\ d\sigma'}{\int_0^{\infty} \mathcal{L}(\sigma', \epsilon, L|n)\ d\sigma'} = 0.95,$$

which in this simple case can be even solved analytically.
The generalization to a multiple bin counting experiment is trivial. In this case indeed, for each bin i we have n_i observed and $\mu_i = \epsilon\sigma L s_i + B_i$ expected events, where s_i is the bin value of the signal density distribution s [2] and B_i is the number of background events. We assume - as before - to know exactly s_i and B_i for each bin.
The likelihood associated to the ith bin is the same as in the example before. The binned likelihood is thus:

$$\mathcal{L}(\sigma, \epsilon, L, s_i, b_i|n) = \prod_{i=1}^{N} \frac{\mu_i^n}{n_i!} e^{-\mu_i}$$

where N is the total number of bins. Equipped with this new $\mathcal{L}$ definition we can run through the same steps described in the previous example and eventually find the upper limit σ_{95}.

10.3 Limit Calculation including Systematic Effects

So far we have disregarded any uncertainty on the parameter values. Now we will fold the systematic errors on ϵ, L, s_i and B_i in the limit calculation. Before doing that we need however to address some issues related with these uncertainties:

B_i: One might ask why we did not factorize the luminosity for the expected background and we did not write $B_i = Lb_i$ as we did for the signal. This would certainly be the right thing to do if our background prediction was based on a Monte Carlo calculation (as it actually is for the signal). However one of the strengths of our analysis is that we do not predict the background, *we measure it*, extrapolating the sideband fit of the data dijet mass distribution inside the signal region (sec. 9.7) [3]. Therefore we do not need the luminosity information for the background prediction, we directly have the expected number of background events in our sample. The main consequence of this approach is that the uncertainty on B_i is independent from the error on L because, in the end, B_i does not depend on that measurement at all. In this way we can avoid all the

[2] Hence we have $\sum_{i=1}^{N} s_i = 1$, where N is the number of bins. We need the density distribution for the signal because we have factorized out L.

[3] It is important to stress that this has nothing to do with the fact that we used the data as a background template in the neural network training. As a matter of fact the method used to fit the background is utterly independent from the event selection procedure.

complications associated with the presence of L in both the signal and background predictions. This would lead indeed to correlations between these two terms that are not straightforward to disentangle. In contrast, our *independent* measurement of B_i has as uncertainty only for the errors associated to the fit of the sidebands. We have already shown that this fit is very stable and this is translated in a small error on B_i. Since the error on s_i is much bigger, as we will see in a moment, we will assume in the following calculation a negligible error on B_i. It is certainly irrelevant for the limit assessment.

L: The value of L is affected by the uncertainty on the instantaneous luminosity measurement. It has been extensively studied elsewhere [53] and the quoted systematic error is 6%.

ϵ: We have already discussed the systematic errors of our selection efficiency (sec. 7.4) but as far as the calculation of the limit is concerned it is convenient to break it up in two pieces; one consisting of those systematics that can modify both the shape of the expected signal distribution and the number of events accepted and the other consisting of systematics that affect only the rate of accepted events. We denote the former with a subscript `F` and the latter with a subscript `R`. We can write therefore the efficiency as $\epsilon = \epsilon_F \cdot \epsilon_R$ and define the corresponding uncertainties as $\Delta\epsilon_F$ and $\Delta\epsilon_R$ respectively. In our case it is not difficult to identify which efficiencies are of the first kind, we know they are those associated with the selection criteria involving jets. Using the same notation we used throughout this dissertation we have [4]:

$$\begin{cases} \epsilon_R = \epsilon_{trg} \cdot \epsilon_{zvtx} \cdot A_{kin} \cdot \epsilon_{ph} = 0.144 & \texttt{with } \Delta\epsilon_R = 10.0\% \\ \epsilon_F = \epsilon_{jets} \cdot \epsilon_{NN} = 0.32 & \texttt{with } \Delta\epsilon_F = 28.8\% \end{cases}$$

As we will see in a moment this distinction is essential for a proper calculation of the limit.

s_i: Since s is a density distribution, it represents the shape of the predicted signal. For this reason it is not affected by the uncertainties of kind `R` but it is sensitive to those of kind `F`. The latter induces an uncertainty Δs_i on the s_i number of signal events in the ith bin. Δs_i can be both positive or negative according to the modification of the distribution shape that can either enrich or deplete of events the ith bin. The correlation between Δs_i and $\Delta\epsilon_F$ requires a special care in handling these two uncertainties.

Again, to illustrate the method we start with a simplified example where we assume $\Delta\epsilon_F = 0$ and, thus, $\Delta s_i = 0$. The limit calculation is done in the same way we showed in the previous paragraph, the only difference is the choice of

[4] For the trigger efficiency ϵ_{trg} we used the weighted average of the efficiencies for the two versions of the trigger $\epsilon_{trg} = (\epsilon_{trg}^{DPS} L^{DPS} + \epsilon_{trg}^{dj12} L^{dj12})/L = 0.781$.

the prior density $\pi(\sigma, \epsilon, L, s_i, b_i)$. Now it must incorporate the uncertainties [5] for the parameters ϵ and L (s_i and b_i are assumed exactly known). Since they are independent we can sum this uncertainties in quadrature. Thus we can define the parameter $\alpha = \epsilon L$ with an uncertainty $\Delta\alpha$ and the posterior density becomes :

$$p(\sigma|n) = \int \mathcal{L}(\sigma, \alpha'|n)\pi(\alpha')\pi(\sigma)\ d\alpha'$$

where we omitted the posterior normalization factor. Since the value of α and its uncertainty are known, it seems reasonable to use for $\pi(\alpha)$ a gaussian with mean α and sigma $\Delta\alpha$. However this choice would lead to serious integration problems, in particular if a constant prior is used for σ, in which case the integral is divergent. A simple way to avoid this problem is to use as a prior for α a gamma distribution [6]. Now, even with a flat prior for σ, the posterior $p(\sigma|n)$ is well defined and the limit calculation can be successfully completed.
Let's now consider the general case when we have an error on the shape of the signal density function as well. The problem is - as before - to find a form for the prior $\pi(\sigma, \epsilon, L, s_i, b_i)$. The b_i are assumed to be known exactly and thus we can use a Dirac delta distribution for them. The luminosity L comes with an uncertainty ΔL and it can be treated with a Gamma distribution, as we did before for α. Setting $\pi(\epsilon, s_i)$ instead entails some more work.
First of all we separate ϵ in the `R` and `F` components. Since ϵ_R is independent from s_i, its prior is the usual Gamma distribution. To take into account the shape change of the signal density we rewrite the expected number of signal events as $(s_i + t\Delta s_i)$ where t - the same for all bins - parametrizes the ignorance on s_i. In other words the signal density s - that was fixed so far - is now depending on the parameter t that gives the "size" of the signal shape uncertainty. For $t = 0$, of course, we have the predicted signal density. In general, however, $\sum_{i=1}^{N}(s_i + t\Delta s_i) \neq 1$ because - as we have seen when we calculated ϵ_{NN} and ϵ_{jets} - a change in shape results in a change in the efficiency, implying the total number of signal events (inside the mass window) not to be constant anymore [7]. This means also that the error $\Delta\epsilon_F$ is now folded in the signal prediction and therefore we do not need to take into account the error on ϵ_F in its prior (*i.e.* we can use a Delta distribution instead of a gamma distribution). In fact, from how we calculated $\Delta\epsilon_F$ and Δs_i we have:

$$\sum_{i=1}^{N}(s_i + \Delta s_i) = 1 + \frac{\Delta\epsilon_F}{\epsilon_F} \quad (t = 1)$$

[5] This step of the limit calculation is often called *likelihood smearing*. In the fully Bayesian approach we adopted instead the presence of uncertainties are naturally included in the prior density functions and they are conceptually separated by the actual form of the likelihood.

[6] The gamma distribution is a two parameter function defined as $\gamma(x|a, b) = \frac{x^{a-1}e^{-x/b}}{\Gamma(a)b^a}$ and with mean $\mu = ab$ and width $\sigma = \sqrt{a}b$. In our case we can determine a and b solving the two equations $\alpha = \mu = ab$ and $\Delta\alpha = \sigma = \sqrt{a}b$.

[7] Of course $\sum_{i=1}^{N} s_i = 1$ still holds.

This is a nice and compact way to appreciate the interplay between t, Δs_i and $\Delta\epsilon_F$. Finally, since we have introduced a new parameter t, we need a prior for it. We defined the Δs_i in such a way that the prior for t should be gaussian with mean 0 and width 1.
Outfitted with all the necessary ingredients, we can now calculate the posterior density (we omit the posterior normalization factor):

$$p(\sigma|n) = \iiint \mathcal{L}(\sigma, \epsilon_F, s_i, \Delta s_i, b_i, \epsilon'_R, L', t')\gamma(\epsilon'_R|\epsilon_R, \Delta\epsilon_R)\gamma(L'|L, \Delta L)\mathcal{G}(t'|0,1)\, d\epsilon'_R\, dL'\, dt'$$

where $\gamma(x|\mu,\sigma)$ and $\mathcal{G}(x|\mu,\sigma)$ are respectively a Gamma and a gaussian distribution with both mean μ and width σ. The expression for the likelihood is:

$$\mathcal{L}(\sigma, \epsilon_F, s_i, b_i, \epsilon_R, L, t) = \prod_{i=1}^{N} \frac{\mu_i^n}{n_i!} e^{-\mu_i}$$

with $\mu_i = \sigma\epsilon_F\epsilon_R L(s_i + t\Delta s_i) + b_i$. With the posterior density in hand it is now straightforward to derive the 95% C.L. upper limit cross section σ_{95}.

10.4 Limit Calculation Results

After having acquired all the basic knowledge on how to set an upper limit, we now apply it to calculate the 95% CL upper limit on the $W(Z)\gamma \to \gamma q\bar{q}$ cross section. The information we need to derive the posterior density $p(\sigma|n)$

Likelihood Parameters	
Parameter	*Value*
Mass Window	60-120
Events Expected (n)	11692
Bin Size	4 GeV
Number of Bins (N)	15
ϵ_F	0.32
$\lvert\Delta\epsilon_F\rvert$	0.092
ϵ_R	0.144
$\lvert\Delta\epsilon_R\rvert$	0.014
L	184
$\lvert\Delta L\rvert$	10.9

Table 10.1: *Parameters used in the likelihood calculation for the determination of the $W(Z)\gamma \to \gamma q\bar{q}$ cross section limit. The N values of s_i, Δs_i and b_i are omitted.*

is summarized in Table 10.1. We omitted all the values of s_i, Δs_i and b_i that are - it is important to remember - fixed. The integration of the likelihood was done numerically using the following procedure:

1) We generated random numbers from gaussian and gamma distributions to get the values for (ϵ'_R, L', t').

2) Using the known values of ϵ_F, s_i, Δs_i and b_i and the three numbers (ϵ'_R, L', t') from 1) we computed $\mathcal{L}(\sigma, \epsilon_F, s_i, \Delta s_i, b_i, \epsilon'_R, L', t')$ for a given σ.

3) We repeated 1) and 2) (with the same σ) 10^4 times and we calculated the average outcome for $\mathcal{L}$, $\bar{\mathcal{L}}(\sigma)$.

4) Repeating 3) for different σ we built the $\bar{\mathcal{L}}(\sigma)$ function point by point.

5) We normalized $\bar{\mathcal{L}}(\sigma)$.

The normalized $\bar{\mathcal{L}}(\sigma)$ distribution is exactly what we want, the posterior density function $p(\sigma|n)$. To find the upper limit we then solved the equation:

$$\int_0^{\sigma_{95}} p(\sigma|n) = 0.95$$

The result was:

$$\begin{aligned}\sigma_{W(Z)\gamma} &= \sigma(p\bar{p} \to W\gamma) \times \mathfrak{B}(W \to q\bar{q}) + \sigma(p\bar{p} \to Z\gamma) \times \mathfrak{B}(Z \to q\bar{q}) \\ &< 54 \text{ pb} \quad \text{at } 95\% \text{ CL}\end{aligned}$$

compared to a Standard Model expectation of 20.50 ± 2.53 pb.

10.5 Future Prospects

Although we were not able to extract any signal peak with 184 pb^{-1} of data at our disposal, the combination of a dedicated high efficient trigger with an optimized advanced offline selection offers a real opportunity to find the W/Z hadronic peak with the data that CDF has been collecting in Run II. With the conservative assumption that the sensitivity of this analysis will not be further improved in the future, we can see from fig. 10.1 that the prospect for a possible evidence of a W/Z signal is really good. If the delivered luminosity schedule is respected, a W/Z signal could be identified really soon, with the peak becoming more and more solid with more data to come. A robust W/Z dijet peak (770 signal events per fb are expected in the dijet mass window 72-110 GeV) can highly benefit the calibration of the jet energy scale and the success of dijet mass resolution improvement studies. Moreover the advanced selection we have developed in this dissertation - which was shown to be essential for the big enhancement in sensitivity - can undoubtedly be applied, with some suitable modifications, to other studies that need to discriminate jets coming from the boson decays and those coming from the QCD background. Examples of such studies are the search of heavy diboson production WW and WZ in the semileptonic decays and, of course, the production of the Higgs boson decaying into two b jets.

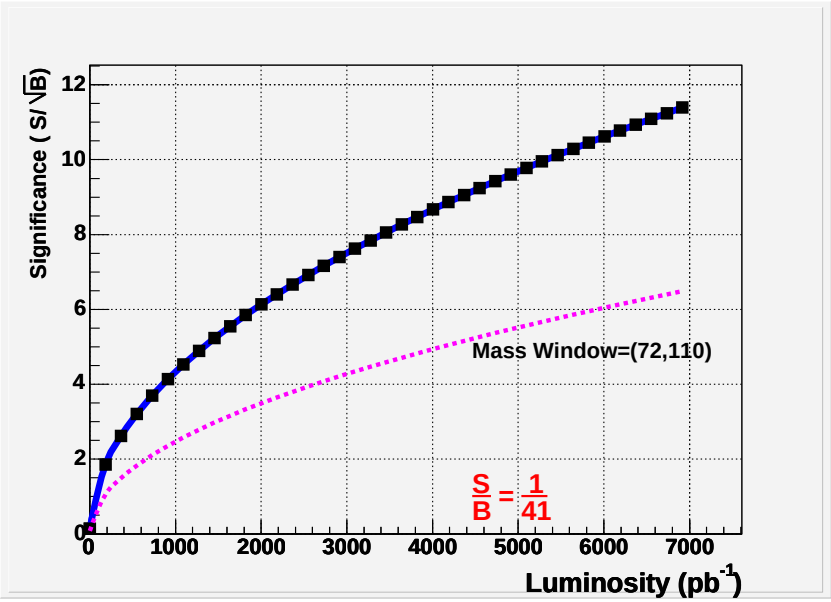


Figure 10.1: *Significance of the W/Z signal as a function of the integrated luminosity after the advanced NN selection described in this dissertation (blue line). As a comparison the same quantity is shown in the case of a standard selection only (dash purple line). The big enhancement in sensitivity due to the NN selection is evident.*

Chapter 11

Summary and Conclusions

A study of the process $p\bar{p} \rightarrow W(Z)\gamma \rightarrow \gamma q\bar{q}$ was presented with the aim of identifying the hadronic decays of the intermediate vector bosons. An advanced selection technique using neural networks was developed for this purpose, which was shown to be very effective for discriminating the signal from the huge QCD background. However, since no evidence of a W/Z signal peak was found in the data, a 95% confidence level upper limit was set.
This analysis is based on the selection of low P_T photons produced in association with at least two jets. To increase the sensitivity of this study, part of the selection requirements were included in a new dedicated trigger implemented in the CDF data acquisition system in August 2003. The $W(Z)\gamma \rightarrow \gamma q\bar{q}$ event candidates were required to have one tightly selected low P_T photon (P_T^{γ} >12 GeV) and two high energy jets (E_T >15 GeV). A neural network advanced selection was applied to further improve the purity of the sample. Finally, a photon background subtraction was performed to eliminate very collimated electromagnetic jets misleadingly identified as isolated photons.
The data sample was from approximately 184 pb^{-1} of proton-antiproton collisions at $\sqrt{s}$ = 1.96 TeV. The efficiency of the event selection was studied using a Monte Carlo simulation of the $W(Z)\gamma \rightarrow \gamma q\bar{q}$ process based on Standard Model predictions. A Monte Carlo estimation of the background was not necessary as it was measured directly from the data. A limit on the $\sigma(p\bar{p} \rightarrow W\gamma) \times \mathfrak{B}(W \rightarrow q\bar{q}) + \sigma(p\bar{p} \rightarrow Z\gamma) \times \mathfrak{B}(Z \rightarrow q\bar{q})$ was set, as no evidence of a γ+W/Z resonance was found.

11.1 Event Selection

Since the existing high P_T inclusive photon triggers did not provide enough sensitivity for the $W(Z)\gamma \rightarrow \gamma q\bar{q}$ search, a new trigger was designed tailored to the analysis needs. The P_T threshold of the photon was lowered to 12 GeV and the requirements of a jet with $|\eta| < 1.78$ as well as certain amount of $\sum E_T$ in the event were added at Level 2. The photon requirements at Level 3 were

also tightened with respect to the other photon triggers. The average trigger efficiency for the identification of $W(Z)\gamma \to \gamma q\bar{q}$ events with $P_T^\gamma > 12$ GeV was measured to be 75.8% (for the last version of the trigger).
A basic set of $W(Z)\gamma \to \gamma q\bar{q}$ candidates was then selected requiring two jets with E_T >15 GeV and setting a veto on other jets with E_T <10 GeV to suppress the contribution from the final state radiation.

11.2 Simulation of $W(Z)\gamma \to \gamma q\bar{q}$ events

A sample of $W(Z)\gamma \to \gamma q\bar{q}$ simulated events was generated using the PYTHIA Monte Carlo generator followed by the CDF standard detector simulation (version `5.3.3`). The bremsstrahlung contribution $p\bar{p} \to W(Z) \to q\bar{q}\gamma$ (usually included in the $W(Z)\gamma$ process) was considered as a background in this analysis and therefore excluded from the signal event generation. The effect of this decision was checked against the prediction of a matrix element Monte Carlo generator (MadGraph).
The prediction for the $\sigma(p\bar{p} \to W\gamma) \times \mathfrak{B}(W \to q\bar{q}) + \sigma(p\bar{p} \to Z\gamma) \times \mathfrak{B}(Z \to q\bar{q})$ was found to be 20.50 ± 2.53 pb for P_T^γ >10 GeV and $|\eta^\gamma| < 1.2$. The total efficiency of this selection was estimated to be 6.2%, which could result in 220 $W(Z)\gamma \to \gamma q\bar{q}$ events expected to be present in our sample (before the neural network selection).

11.3 Neural Network Selection

To further improve the sensitivity of our search, a neural network was built in order to exploit all the possible differences between the signal and the huge QCD $\gamma + 2$ jets background. Several input variables were scrutinized and a method to select only those with the overall highest discrimination power was devised. The effect of the multiple interactions was carefully appraised. Eventually, a cut based on the output of the neural network was applied resulting in an efficiency of 70%. The neural network based selection improved the S/B ratio from 1/333 to 1/71. A study of the signal mass window that maximizes the significance was also carried out The best choice was found to be the 72-110 GeV region, with S/B=1/41 and a significance of 1.86.

11.4 $W(Z)\gamma \to \gamma q\bar{q}$ Peak Search and Cross Section Limit

The dijet mass distribution of the final selected events was used to determine the background spectrum. The two sidebands, on the low and high side of the signal region, were fitted with an exponential and then extrapolated inside the signal region. The use of two legs for the fit highly improved the stability and precision of the background prediction. No event excess, with respect to the

background expectation, was observed. A limit was calculated using a fully Bayesian approach and all the effects of the efficiency systematic uncertainty were taken into account. The 95% confindence level upper limit for the $p\bar{p} \to W(Z)\gamma$ process, with the W/Z bosons decaying hadronically, was set to 54 pb, consistent with the Standard Model prediction of 20.5 pb.

Appendix A

Fermilab Complex

The Tevatron is a proton-antiproton synchrotron accelerator located in Batavia, Illinois, at the Fermi National Accelerator Laboratory (Fermilab). It is currently the source of the highest energy $p\bar{p}$ collisions at the center-of-mass of $\sqrt{s} = 1.96$ GeV. The final $p\bar{p}$ collisions are achieved after several intermediate stages involving preparing the protons and antiprotons, injecting them into the Tevatron, accelerating the two beams up to the energy of 960 GeV and colliding them at the interaction regions. In the following section we briefly describe each of these steps (fig. A.1). A more detailed description of the Tevatron complex can be found in the reference [54].

A.1 From H^- to $p\bar{p}$ collisions

The whole process begins in the Cockcroft-Walton pre-accelerator where H^- ions are created and accelerated to 750 keV of kinetic energy. Then they enter a 150 m long Linear Accelerator (linac) that brings their energy up to 400 MeV. After the Linac they are transferred into the Booster, a circular synchrotron 74.5 m in diameter. A carbon foil traps the electrons attached to the ions, leaving base protons that are accelerated to 8 GeV with RF and sent into the Main Injector.
The Main Injector is the largest upgrade to the Tevatron facility for the Run 2 experiment. It replaced the Main Ring, a 400 GeV proton synchrotron situated in the Tevatron tunnel, built during the early 1970's and later adapted to inject particles into the Tevatron. The Main Injector, a synchrotron of about 3 km in circumference, has four different functions: (1) it accelerates protons to 120 GeV and deliver them to the antiproton production target, (2) it accelerates protons and antiprotons to 150 GeV, (3) it decelerates the 150 GeV antiprotons from the Tevatron and transfers them to the Recycler, and (4) injects protons and antiprotons into the Tevatron. In collider mode, after accelerating the protons from the booster (collected in bunches of 6×10^{10} protons) to 150 GeV, it coalesces them into a single bunch of 27×10^{10} protons and injects it into the

Tevatron and accelerates it to 960 GeV. This represents one MI cycle, thirty-six of them are needed to fill the Tevatron.

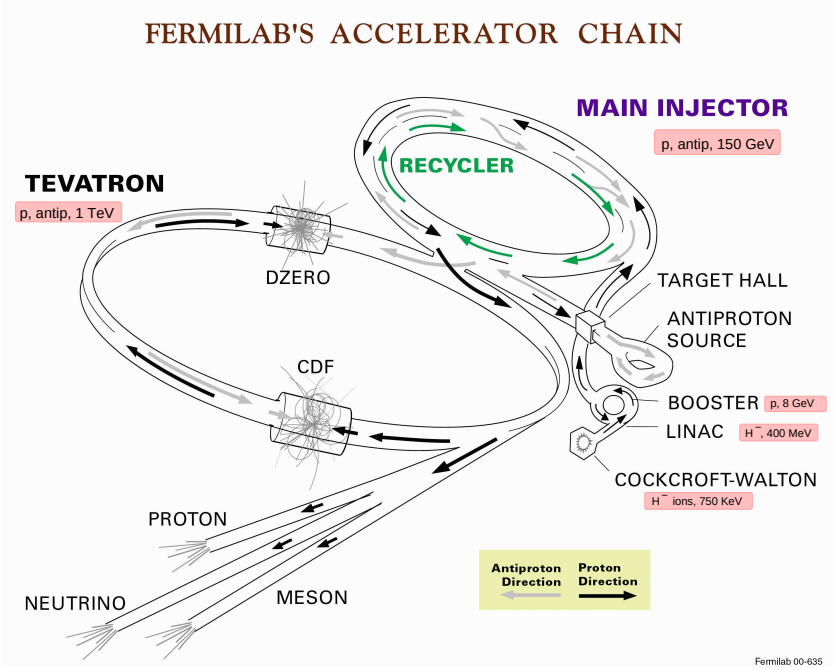


Figure A.1: *The Fermilab accelerator chain*

Antiproton Production

To produce antiprotons, 120 GeV protons from the MI are directed into a nickel target. In the collision, about 20 $\bar{p}$ are produced per one million protons. Then the $\bar{p}$ are focused by a lithium lens and separated from other particle species by a pulsed magnet. The selected particles are fed into the Debuncher where the process of stochastic cooling reduces the energy spread while maintaining the center value of 8 GeV. In this process the $\bar{p}$ beam is transformed into a continuous beam to allow the cooling. The $\bar{p}$ are then transferred and stacked into the Accumulator. The process continues for hours until the maximum Accumulator intensity is reached (it takes about 12-24 hours). At this point the $\bar{p}$ beam is re-bunched and injected into the MI, accelerated to 150 GeV and sent to the Tevatron. Each $\bar{p}$ bunch contains about 33×10^9 particles. The precious antiprotons left at the end of a collider store (the period of time when the colliding beams are retained in the machine) are recovered and re-used thanks to the new Recycler Ring, installed in the Main Injector enclosure. About 75% of the original antiprotons survive and instead to be dumped, they

are re-cooled and sent back to the MI to be ready for the next store. The benefit of this recycling procedure is estimated to be a factor 2 in luminosity.

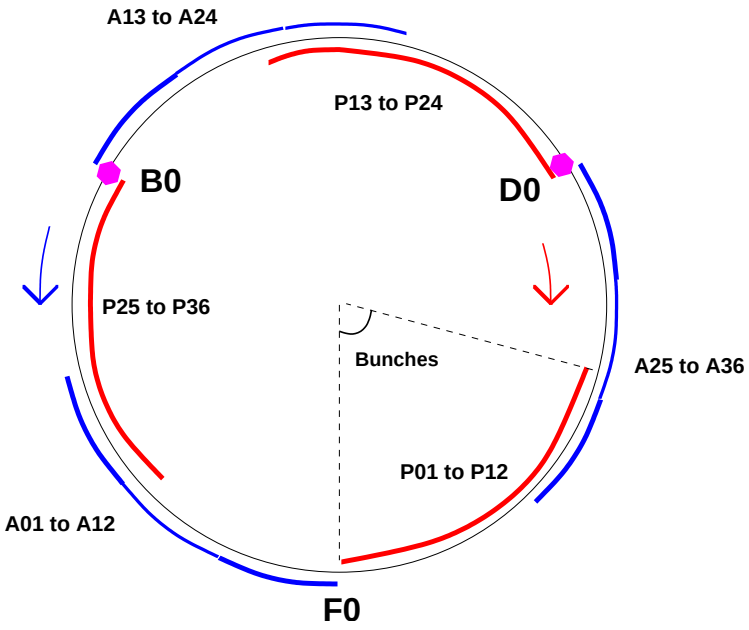


Figure A.2: *Schematic representation of the Tevatron bunch orientation.*

The Tevatron

The final stage of the acceleration is the 6 km long Tevatron ring. The antiprotons are injected after the protons. The two beams are separated and a pair of non-intersecting closed helical orbits are created. In this way the beam-beam electromagnetic interactions are reduced. The beam configuration and injection scheme are illustrated in fig. A.2. Both protons and antiprotons circulate in three trains of 12 bunches each spaced 396 ns, and the three trains are separated by about 2.6 μs ("abort gap"). Since the proton injection happens first, the antiprotons need to be injected during the time that the proton abort gap passes. Once the Tevatron loading is complete, beams are accelerated to the maximum energy and collisions are initiated. Among the 72 regions the bunch crossings occur, in two of them, at B0 and D0, a special afford is made to maximize the chance of a proton-antiproton collision, focusing the beam with quadrupole magnets. This affects the luminosity $\mathcal{L}$:

$$\mathcal{L} = \frac{N_B N_p N_{\bar{p}} f}{2\pi(\sigma_{\bar{p}}^2 + \sigma_p^2)}$$

where N_B is the number of bunches, $N_{(\bar{p})p|}$ is the number of the (anti)protons in a bunch, f is the bunch revolution frequency, and $\sigma_{(\bar{p})p}$ is the effective width of the (anti)proton beam. The *integrated* luminosity, defined as $L = \int \mathcal{L} dt$, is more relevant to physics process in $p\bar{p}$ collisions[1]. In fig. A.3 and fig. A.4 the evolution of the integrated and instantaneous luminosity in Run II are shown.

[1]It is conveniently measured in inverse barn (1 barn=10^{24} cm^2)

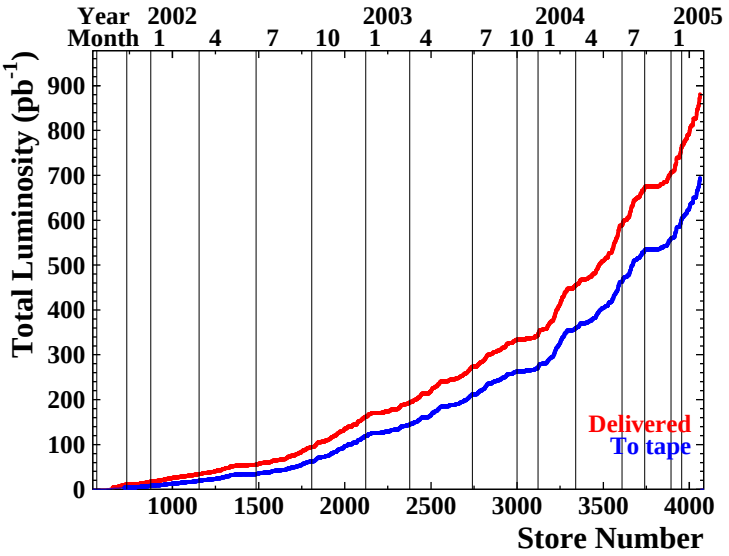


Figure A.3: *Total integrated luminosity delivered and recorded.*

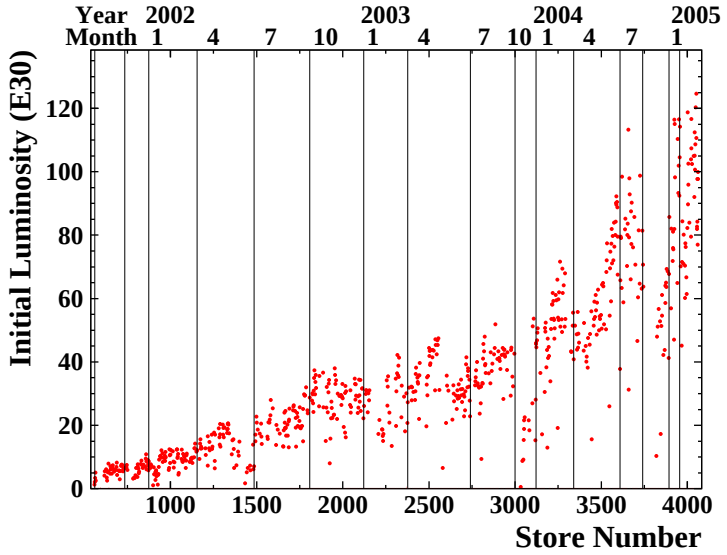


Figure A.4: *Initial luminosity per Tevatron store.*

Bibliography

[1] S. L. Glashow, Nucl. Phys. **22**, 579 (1961)
S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967)

[2] D.J.Gross and F.Wilczek,
Phys. Rev. Lett **19**, 1264 (1967) ;

S.Weinberg
Phys. Rev. D **8** 3633 (1973) ;

H. Fritzsch, M. Gell-Mann, and H. Leutwyler
Phys. Rev. Lett **30**, 1343 (1973);

[3] J.C.Collins, D.E.Soper and G.Sterman
Factorization of hard processes in QCD
in "Perturbative QCD", World Scientific - Singapore

[4] UA2 Collaboration, M. Banner *et al.*, Phys. Lett. B **122** 476 (1983)
UA1 Collaboration, G. Arnison *et al.*, Phys. Lett. B **122** 103 (1983) ;

[5] P. W. Higgs, Phys. Rev. Lett. 12 (1964) 132;

[6] LEP Collaborations, arXiv:hep-ex/0412015;

[7] L. Babukhadia *et al.*
CDF and D0 Working Group Members,
FERMILAB-PUB-03-320-E ;

[8] UA2 Collaboration, J, Alitti *et al.*, Z. Phys. **C49** 17 (1991);

[9] A. Bocci *et al.*, CDF/ANAL/JET/CDFR/4681,
A. Bocci *et al.*, CDF/ANAL/JET/CDFR/5179,
A. Bocci *et al.*, CDF/ANAL/JET/CDFR/5342,
A. Bocci *et al.*, CDF/ANAL/JET/CDFR/5416,
A. Bocci *et al.*, CDF/ANAL/JET/CDFR/5507;

[10] R. Wilkinson *et al.*, CDF/ANAL/JET/CDFR/4191;

[11] T. Dorigo *et al.*, CDF/ANAL/ELECTROWEAK/CDFR/4577;

[12] J.F. Arguin, *et al.*, CDF/PHYS/TOP/GROUP/7505;

[13] M. Brozovic *et al.*, CDF//JET//7389;

[14] CDFII Technical Design Report, FERMILAB-96-390-E;

[15] T. Affolder, *et al.*, Nucl. Phys. **A52**, 249 (2004);

[16] L. Balka, *et al.*, Nucl. Phys. **A267**, 272 (1988);

[17] S. Bertolucci, *et al.*, Nucl. Phys. **A267**, 301 (1988);

[18] M. Albrow *et al.*, Nucl. Phys. **A480**, 524 (2002);

[19] S. Kuhlmann, *et al.*, Nucl. Phys. **A518**, 39 (2004);

[20] D. Acosta,*et al.*, Nucl. Phys. **A518**, 605 (2004);

[21] D. Acosta,*et al.*, Nucl. Phys. **A461**, 540 (2001);

[22] K. Goulianos, *et al.*, Nucl. Phys. **A519**, 42 (2004);

[23] A. Mayer, FERMILAB-01/243-E;

[24] H. Ray, CDF/DOC/TRIGGER/PUBLIC/5860;

[25] G.C. Blazey , *et al.*, arXiv:hep-ex/0005012;

[26] A. Bhatti, *et al.*, CDF/PHYS/JET//7358;

[27] S. Kuhlmann, *et al.*, CDF/PHYS/JET/CDFR/1963;

[28] S. Kuhlmann, CDF/ANAL/JET/CDFR/1392;

[29] R. M. Harris, CDF/ANAL/JET/CDFR/2318;

[30] H. Gerberich *et al.*, CDF/ANAL/EXOTIC/CDFR/6368;

[31] M. Kim *et al.*, CDF/PUB/CDF/PUBLIC/6101;

[32] U. Baur *et al.*, Phys. Rev. D **50**, 1917 (1994);

[33] H.-U. Bengtsson and T.Sjostrand Comp. Phys. Comm. **46** 43 (1987);
T.Sjostrand, Comp. Phys. Comm. **82** 74 (1994);

[34] G.Marchesini and B.R.Webber, Nucl. Phys. B **310** 461 (1988);
G.Marchesini, *et al.*, Comp. Phys. Comm. **71** 15 (1992;

[35] F. Maltoni and T. Stelzer, JHEP **0302**, 027 (2003;

[36] CTEQ Collaboration, J. Huston *et al.*, Eur. Phys. J. C **12**, 375 (2000);

[37] D. Benjamin *et al.*, CDF/PHYS/ELECTROWEAK/CDFR/6601;

[38] U. Baur *et al.*, Phys. Rev. D **48**, 5140 (1993);
U. Baur *et al.*, Phys. Rev. D **57**, 2823 (1998);

[39] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, Eur. Phys. J. C **23**, 73 (2002);

[40] D. Acosta *et al.*, The CDF Collaboration, Phys. Rev. Lett. **94**, 041803 (2005);

[41] W.K. Sakumoto *et al.*, CDF/ANAL/ELECTROWEAK/CDFR/6331;

[42] R. Culbertson *et al.*, CDF/ANAL/EXOTIC/CDFR/7024;

[43] Y. Liu *et al.*, CDF/PHYS/CDF/PUBLIC/6312;

[44] A. Bocci *et al.*, CDF/ANAL/JET/CDFR/4259;

[45] A. Bocci *et al.*, CDF/ANAL/JET/CDFR/6042;

[46] H. Baer *et al.*, Phys. Rev. D **42**, 61 (1990);
S. Catani *et al.*, arXiv:hep-ph/0204023;

[47] C.M. Bishop,
Neural Networks for Pattern Recognition,
Clarendon, Oxford, 1998;

[48] R.A. Fisher, Annals of Eugenics, 7:179 (1936);

[49] A ROOT Interface to JETNET, CDF/ANAL/TOP//5434;

[50] C. Neu *et al.*, CDF/ANAL/EXOTIC/CDFR/6309;

[51] P. Azzi *et al.*, CDF/ANAL/TOP/CDFR/6808;

[52] L. Demortier, CDF/MEMO/STATISTICS/PUBLIC/5928;

[53] S. Jindariani *et al.*, CDF/ANAL/CDF//7446;

[54] Fermilab Rookie Books,
http://www-bdnew.fnal.gov/operations/rookie_books/rbooks.html;